

# 企业数字档案馆建设 理论与实践

蔡盈芳 著

電子工業出版社

Publishing House of Electronics Industry

北京 · BEIJING

## 内 容 简 介

本书从信息论的角度,首先介绍了企业数字档案馆层次架构,各部分的组成;然后详细介绍了企业数字档案馆建设的主要内容,包括软硬件建设、数据资源建设、软件开发、管理机制建设、标准规范体系建设、安全保密体系建设,以及建设思路。本书还就应用软件工程理论,详细论述了企业数字档案馆建设的过程、注意事项,以及成果评价等,尤其对电子档案管理系统的开发与实施,进行了系统论述。为了进一步提供实践依据,本书分别从集团型企业总部、基层企业等不同角度进行选材,介绍了四个企业的数字档案馆建设案例。

本书适合企业档案工作者、各级档案行政管理人员、从事档案管理信息化业务的服务商、高校相关专业师生,以及档案科研机构的专业人员学习。

未经许可,不得以任何方式复制或抄袭本书之部分或全部内容。  
版权所有,侵权必究。

## 图书在版编目(CIP)数据

企业数字档案馆建设理论与实践 / 蔡盈芳著. —北京:电子工业出版社, 2018.9  
ISBN 978-7-121-35128-0

I. ①企… II. ①蔡… III. ①数字技术—应用—档案馆—研究 IV. ①G270.7

中国版本图书馆 CIP 数据核字(2018)第 221925 号

策划编辑:许存权

责任编辑:许存权 特约编辑:谢忠玉 等

印 刷:

装 订:

出版发行:电子工业出版社

北京市海淀区万寿路 173 信箱 邮编 100036

开 本:720×1 000 1/16 印张:17 字数:326 千字

版 次:2018 年 9 月第 1 版

印 次:2018 年 9 月第 1 次印刷

定 价:89.00 元

凡所购买电子工业出版社图书有缺损问题,请向购买书店调换。若书店售缺,请与本社发行部联系,联系及邮购电话:(010) 88254888, 88258888。

质量投诉请发邮件至 [zlt@phei.com.cn](mailto:zlt@phei.com.cn), 盗版侵权举报请发邮件至 [dbqq@phei.com.cn](mailto:dbqq@phei.com.cn)。

本书咨询联系方式:(010) 88254484, [xucq@phei.com.cn](mailto:xucq@phei.com.cn)。



信息技术广泛深入的应用是这个时代的一大特征，在这个大背景下，企业形成待归档的电子文件越来越多，并呈现几何级数增长态势，电子文件归档显得更加迫切。随着时代的进步，企业科研、生产、经营、管理各项业务对档案工作提出了更加高效的需求。为适应这些需求，企业档案工作数字化转型成为必然。在国外，档案工作的数字化转型成为推动档案事业发展的主要动力，实现数字化转型正成为国外档案工作的主要内容。《全国档案事业发展“十三五”规划纲要》提出，要建成一批具有国际先进水平的企业数字档案馆。近几年，在数字化转型需求的驱动下，在有关规划政策的指引下，企业建设数字档案馆的积极性进一步提高，正成为企业档案工作提质增效的有效措施。企业数字档案馆建设是一项系统工程，需要有多个学科知识支撑，并综合运用，融会贯通。本书融合了档案管理理论和信息技术理论，结合当前企业数字档案馆建设的实际需求，阐述了企业数字档案馆建设的内容和过程，既涉及企业数字档案馆“是什么”，又涉及企业数字档案馆“怎么建”，囊括了企业数字档案馆建设的全过程。

本书第一章阐述了企业数字档案馆建设的理论框架，构建企业数字档案馆建设的主体架构；第二、三章主要讨论企业数字档案馆的建设过程；第四至九章按照企业数字档案馆主体架构的分布，逐个详细讨论其建设内容和方法，由于电子档案管理体系的建设比较复杂，内容较多，故用了第七、八章两章进行叙述；管理机制内容较少，未独立成章，有关内容穿插在有关章中；第十、十一章是补充和扩展内容，就企业数字档案馆的评价、几项建设中如何应用新技术作了讨论，供广大档案工作者借鉴参考；第十二章是四个案例，从不同角度验证前述章节的内容，在案例选择上，力求具有代表性，其中某央企集团公司、东风汽车公司案例主要供有下属单位的企业总部档案人员参考，国泰君安、本钢集团公司案例可供其他企业档案人员参考。每一章节的内容安排均按照先理论再实践的顺序展开，即先叙述信息系统建设的一般原理，再叙述数字档案馆建设的相应理论，最后根据相应的理论叙述建设数字档案馆的具体实践和方法。尽管有些信息系统建设理论对有信息技术知识（或有一定信息技术技能）的档案人员来说是基础性知识，但考虑到本书是面向大多数档案工作者，有了解信息系统建设理论和基础知识的需求，且这些理论和基础知识有利于后续内容的理解，因此，将这些内容也作为本书的一部分。

本书案例的有关内容分别由东风汽车集团有限公司办公室杨耀红经理、国泰君

安证券股份有限公司档案中心刘俊总监、本钢集团有限公司档案馆于胜连馆长提供，特此致谢！

本书适合企业开展数字档案馆建设的有关人员参考，也可供从事企业档案业务监督指导的档案工作者参考；本书的有关内容还可供电子档案管理系统软件开发人员参考，以便更好地了解企业数字档案馆建设的需求，更好地与企业档案人员沟通。考虑到企业数字档案室建设的内容与数字档案馆建设的内容基本相同，为方便叙述，本书书名及定位均为企业数字档案馆，但同样适合企业开展数字档案室建设时借鉴。

因时间紧迫，加上水平有限，书中错漏和不足之处，恳请读者批评指正。

蔡盈芳

2018年8月于北京



<b>第一章 企业数字档案馆概述</b> .....	(1)
第一节 企业数字档案馆建设背景 .....	(1)
一、建设数字档案馆是国家经济发展的必然要求 .....	(1)
二、建设数字档案馆是档案事业发展的客观要求 .....	(1)
三、建设数字档案馆是企业发展的需求 .....	(2)
第二节 数字档案馆定义、特征及发展历程 .....	(3)
一、数字档案馆定义 .....	(3)
二、数字档案馆基本特征 .....	(4)
三、数字档案馆发展历程 .....	(5)
四、数字档案馆在我国企业的应用 .....	(7)
第三节 企业数字档案馆总体架构 .....	(8)
一、研究企业数字档案馆总体架构的意义 .....	(8)
二、企业数字档案馆总体架构的层次模型及其功能 .....	(9)
三、数字档案馆总体架构层次模型的启示 .....	(11)
四、企业数字档案馆建设内容 .....	(12)
第四节 企业数字档案馆建设步骤 .....	(17)
一、项目规划 .....	(18)
二、项目立项 .....	(18)
三、项目实施 .....	(18)
四、项目验收 .....	(19)
五、项目维护 .....	(19)
第五节 企业建设数字档案馆的效益 .....	(19)
一、顺应档案工作发展的必然要求 .....	(19)
二、助力企业信息化建设 .....	(20)
三、实现档案事业的平衡协调发展 .....	(21)
四、解决当前企业档案工作面临的困难与问题 .....	(22)
五、实现企业档案提质增效 .....	(22)
第六节 数字档案馆与档案信息化阶段理论 .....	(24)
一、诺兰阶段模型及其指导作用 .....	(24)

二、档案信息化阶段的划分 .....	(25)
三、档案信息化各阶段特征模型 .....	(26)
四、档案信息化发展阶段划分的意义 .....	(27)
<b>第二章 企业数字档案馆建设规划 .....</b>	<b>(29)</b>
第一节 数字档案馆规划原理 .....	(29)
一、企业信息系统规划原理 .....	(29)
二、档案信息化规划步骤 .....	(31)
第二节 数字档案馆规划的主要内容 .....	(32)
一、现状分析 .....	(32)
二、环境分析 .....	(33)
三、标杆分析 .....	(34)
四、SWOT 分析 .....	(35)
五、企业档案事业战略和企业信息化战略继承 .....	(35)
六、目标确定 .....	(36)
七、数字档案馆总体技术架构设计 .....	(36)
八、电子档案管理系统平台设计 .....	(36)
九、确定重点任务 .....	(38)
十、制定规划路线 .....	(38)
十一、规划任务的外包 .....	(39)
第三节 规划文本的形成、评审和项目立项 .....	(40)
一、规划文本的形成 .....	(40)
二、规划文本的评审 .....	(42)
<b>第三章 企业数字档案馆建设实施与运维 .....</b>	<b>(43)</b>
第一节 实施准备 .....	(43)
一、制定实施方案 .....	(43)
二、人才队伍建设 .....	(43)
三、成立项目工作组织 .....	(43)
四、建立项目组运作机制 .....	(45)
第二节 项目实施 .....	(45)
一、项目实施启动 .....	(45)
二、基础业务完善与规范化 .....	(46)
三、制度规范制定 .....	(46)
四、基础设施配备 .....	(46)
五、电子档案管理系统实施 .....	(47)

六、传统载体档案数字化实施 .....	(47)
七、电子文件归档实施 .....	(48)
第三节 项目实施验收 .....	(49)
一、验收内容和目的 .....	(49)
二、验收依据 .....	(49)
三、验收方式 .....	(50)
四、验收意见 .....	(50)
第四节 项目运维 .....	(50)
一、信息系统运行管理的原理 .....	(51)
二、信息系统维护的一般原理 .....	(53)
三、数字档案馆运维工作 .....	(55)
<b>第四章 基础设施建设 .....</b>	<b>(58)</b>
第一节 数字档案馆网络及服务器建设 .....	(58)
一、数字档案馆网络平台架构设计 .....	(58)
二、服务器配置选择 .....	(60)
第二节 信息采集设备 .....	(62)
一、扫描仪 .....	(62)
二、视频采集转换设备 .....	(68)
第三节 存储设备 .....	(71)
一、各类存储设备原理 .....	(71)
二、在线存储 .....	(77)
三、近线存储 .....	(79)
四、离线存储 .....	(80)
五、云存储 .....	(81)
<b>第五章 传统载体档案数字化 .....</b>	<b>(82)</b>
第一节 概述 .....	(82)
一、档案数字化原则 .....	(82)
二、档案数字化的组织 .....	(85)
三、数字化技术 .....	(86)
四、扫描文件存储格式 .....	(87)
第二节 纸质档案数字化 .....	(90)
一、档案出库 .....	(90)
二、数字化前处理 .....	(91)
三、档案扫描 .....	(92)

四、图像处理 .....	(94)
五、数据挂接 .....	(94)
六、数字化成果验收与移交 .....	(95)
七、档案归还入库 .....	(97)
第三节 多媒体档案数字化 .....	(97)
一、录音档案数字化 .....	(97)
二、录像档案数字化 .....	(98)
第四节 档案数字化外包安全管理 .....	(99)
一、外包方档案部门的安全管理工作 .....	(99)
二、承包数字化工作的服务机构的安全管理 .....	(100)
三、数字化场所的安全管理 .....	(101)
四、数字化加工设备、网络环境与数据载体的安全管理 .....	(102)
五、档案实体的安全管理 .....	(103)
六、档案数字化成果移交与设备的安全管理 .....	(103)
<b>第六章 电子文件归档和电子档案管理 .....</b>	<b>(105)</b>
第一节 电子文件管理原则和要求 .....	(105)
一、电子文件管理原则 .....	(106)
二、电子文件管理要求 .....	(108)
三、电子文件管理环节 .....	(114)
第二节 电子文件形成的管理 .....	(115)
一、保证电子文件真实、可靠、安全 .....	(116)
二、按一定格式存储 .....	(116)
三、捕获相应的元数据 .....	(117)
第三节 电子文件归档 .....	(118)
一、电子文件归档范围 .....	(118)
二、归档收集方式 .....	(128)
三、归档鉴定 .....	(131)
四、整理 .....	(133)
五、归档移交 .....	(134)
六、归档接收 .....	(136)
第四节 电子档案存储、保管、利用、处置 .....	(137)
一、电子档案存储 .....	(137)
二、电子档案保管 .....	(138)
三、电子档案利用 .....	(146)
四、电子档案处置 .....	(146)



**第七章 电子档案管理系统软件分析与设计** ..... (147)

第一节 电子档案管理系统软件生命周期.....	(147)
一、软件开发周期阶段 .....	(147)
二、系统生命周期阶段 .....	(147)
第二节 档案管理信息系统需求分析.....	(149)
一、需求获取方法 .....	(149)
二、需求陈述方法 .....	(151)
三、系统分析报告 .....	(154)
第三节 电子档案管理系统需求陈述.....	(156)
一、收集功能 .....	(156)
二、整理功能 .....	(157)
三、保管功能 .....	(159)
四、统计功能 .....	(160)
五、利用功能 .....	(160)
六、系统管理功能 .....	(161)
七、传统载体档案辅助管理功能 .....	(163)
八、其他可选功能 .....	(164)
第四节 模块设计.....	(164)
一、模块或子系统划分的依据 .....	(164)
二、模块划分实践 .....	(165)
第五节 界面及输入/输出设计.....	(168)
一、界面设计 .....	(168)
二、输入/输出设计 .....	(170)

**第八章 电子档案管理系统实施** ..... (171)

第一节 电子档案管理系统实施概述.....	(171)
一、系统实施目标 .....	(171)
二、系统实施技术原则 .....	(171)
三、系统实施的主要内容与步骤 .....	(171)
第二节 软件采购.....	(172)
一、初始需求分析 .....	(172)
二、原型测试 .....	(173)
三、商务谈判 .....	(173)
四、签订合同及付款 .....	(173)
五、客户化开发 .....	(173)

六、软件安装 .....	(174)
第三节 基础数据准备 .....	(174)
一、用户数据准备 .....	(174)
二、数据字典建立 .....	(176)
三、界面设计 .....	(177)
四、报表设计 .....	(177)
五、分类结点树方案设计 .....	(178)
第四节 系统测试 .....	(179)
一、系统测试概述 .....	(179)
二、测试方法 .....	(180)
三、测试类型 .....	(181)
四、黑盒测试用例设计 .....	(182)
五、排错 .....	(186)
六、采购测试 .....	(187)
第五节 人员培训 .....	(189)
一、业务和职能部门人员 .....	(190)
二、档案管理人员 .....	(190)
三、系统维护人员 .....	(191)
第六节 系统切换 .....	(191)
一、直接转换 .....	(191)
二、平行转换 .....	(192)
三、分段平行转换 .....	(193)
第七节 系统评价与验收 .....	(194)
一、系统评价 .....	(194)
二、系统验收 .....	(196)

## 第九章 标准规范与安全保密体系建设 .....

第一节 标准规范建设 .....	(198)
一、基础标准规范 .....	(198)
二、管理过程标准规范 .....	(199)
三、管理对象标准规范 .....	(200)
第二节 安全保密体系建设 .....	(201)
一、概述 .....	(201)
二、建设内容 .....	(202)
三、安全与保密重点 .....	(203)

<b>第十章 企业数字档案馆评价</b>	( 204 )
第一节 数字档案馆建设评价问题的提出	( 204 )
一、数字档案馆建设评价	( 204 )
二、基于层次分析法的评价指标体系构建思路	( 205 )
第二节 数字档案馆建设评价总体思路	( 206 )
第三节 企业档案信息化评价指标体系设计	( 207 )
一、评价指标结构的确定	( 207 )
二、权重的确定	( 207 )
<b>第十一章 企业数字档案馆与新技术的融合</b>	( 211 )
第一节 企业数字档案馆与知识管理	( 211 )
一、当前知识管理的突出问题	( 211 )
二、知识管理及其技术	( 212 )
三、企业档案信息特点及知识管理难点分析	( 212 )
四、知识管理技术的运用	( 215 )
第二节 企业数字档案馆与大数据	( 218 )
一、大数据及其特点	( 218 )
二、大数据与档案工作的融合	( 219 )
三、大数据技术在数字档案馆中的应用	( 220 )
第三节 数字档案馆与云服务	( 221 )
一、云服务概述	( 221 )
二、数字档案馆云服务类型	( 223 )
三、数字档案馆云服务机遇和面临的挑战	( 224 )
四、基于云服务的数字档案馆架构及实现	( 224 )
<b>第十二章 数字档案馆建设案例</b>	( 227 )
第一节 某央企集团公司数字档案馆建设案例	( 227 )
一、工作步骤	( 227 )
二、主要内容	( 228 )
三、主要成果	( 230 )
第二节 国泰君安证券公司数字档案室建设案例	( 230 )
一、建设背景	( 231 )
二、建设目标	( 231 )
三、建设内容	( 233 )
四、建设技术方案	( 236 )

五、建设进度 .....	(239)
第三节 东风公司数字档案馆建设项目案例 .....	(240)
一、项目综述 .....	(240)
二、项目一期建设过程 .....	(241)
三、一期项目建设效果 .....	(242)
四、项目建设经验总结 .....	(245)
第四节 本钢集团有限公司数字档案馆建设案例 .....	(248)
一、系统建设概况 .....	(248)
二、系统运行管理体制 .....	(249)
三、科学设计电子文件归档技术方案 .....	(250)
四、建设成效 .....	(253)
五、主要经验 .....	(254)
参考文献 .....	(257)
后记 .....	(259)

# 第一章 企业数字档案馆概述

## 第一节 企业数字档案馆建设背景

### 一、建设数字档案馆是国家经济发展的必然要求

随着我国经济的发展，经济转型升级成为必然要求。互联网+、大数据、人工智能迅速发展，网络强国、数字中国、智慧社会建设初显端倪。互联网与商业结合使电子商务方兴未艾，互联网与工业紧密结合使工业互联网、智能制造迅速发展。电子发票在中国得到广泛应用，信息技术正深入人类生产、生活的各个方面。信息技术的进一步深入应用，使文件数据化、电子化趋势更加明显，仅靠传统手工管理无法完成。建设数字档案馆，实现档案工作对电子文件和电子档案真实、可靠、完整和安全的管理，是国家经济发展对档案工作的必然要求。

### 二、建设数字档案馆是档案事业发展的客观要求

近年来，国家档案局更加重视数字档案馆建设工作，继 2010 年推出《数字档案馆建设指南》后，相继推出《数字档案馆系统测试办法》、《数字档案室建设指南》、《数字档案室建设评价办法》，又于 2017 年颁发了《企业数字档案馆（室）建设指南》。《全国档案事业发展“十三五”规划纲要》提出，要建成一批具有国际先进水平的企业数字档案馆。为此，国家档案局决定开展企业数字档案馆（室）建设试点工作，并于 2017 年年底下发通知并正式启动。2018 年年初确定 35 家单位为企业数字档案馆（室）建设试点单位，2018 年 7 月召开了试点建设工作启动会。

当前开展数字档案馆建设，对档案事业的发展具有特别重要的意义，是解

决当前档案工作突出问题的需要。当前档案工作要解决的突出问题，就是电子文件归档和电子档案管理。通过前期的电子文件归档管理试点工作发现，数字档案馆建设已经成为电子文件归档管理的瓶颈。许多单位都因为没有建设数字档案馆和数字档案室，没有电子档案管理系统，电子文件归档工作无从开展。因此，建设数字档案馆和数字档案室，有利于促进电子文件归档工作。

当前开展数字档案馆建设也是档案供给侧结构性改革的需要。中国共产党第十九次全国代表大会提出，我国社会主要矛盾由“人民日益增长的物质文化需要同落后的社会生产之间的矛盾”转化为“人民日益增长的美好生活需要和不平衡不充分的发展之间的矛盾”。未来一段时间的工作重点就是要解决发展不平衡、不充分的问题。具体到企业档案工作，随着经济发展和企业改革的深化，各方面都对档案工作提出了急迫需求。横向比较来看，全国企业档案工作发展也极不平衡。要解决企业档案工作中的这些问题，一条重要措施就是建设数字档案馆和数字档案室，实现企业档案工作的数字化转型。数字档案馆和数字档案室具有提高档案管理效率，提升档案利用水平等功能。通过它的建设，能有效地解决企业档案工作发展不充分、不平衡的问题。通过数字档案馆和数字档案室建设，可使档案信息化与企业信息化发展水平相适应，与企业档案信息需求相适应，最终使各单位档案信息化水平与国家档案信息化整体水平相适应。

### 三、建设数字档案馆是企业发展的需求

信息化为企业提升效益提供了新的动力。自从信息技术得到应用以来，企业应用信息技术的热情有增无减，从日常管理到技术开发，企业的每一项业务都因信息技术的应用发生了翻天覆地的变化，并促成各个产业的升级换代。随着信息技术在企业广泛深入地应用，对档案工作也提出了新的要求。主要表现在以下三个方面。

第一，企业大量应用信息技术对档案信息管理效率的提高提出了迫切需求。信息技术的应用大幅度提高了企业的效率，在技术研发方面，通过应用信息技术，产品研发周期大为缩短；在产品加工方面，通过应用信息技术，产品加工效率成倍增加，质量进一步提高；在市场销售方面，通过应用信息技术，产品销售模式不断得到创新，市场效率不断得到提高。企业各方面效率的提高，也对企业档案工作的效率提高提出了要求，实践证明，信息化是提高档案工作效

率的有效途径。

第二，信息化条件下的企业要求更方便地获取档案信息。信息技术给企业带来的最显著进步就是信息的获取比以前更加方便，企业信息流动的速度和便捷性超过了以往任何时期。在此大背影下，对作为企业重要信息流一部分的档案信息流在获取的方便与快捷方面提出了要求，在客观上要求档案工作应用信息技术，实现档案工作的信息化。

第三，电子文件的大量出现，客观上要求应用信息技术手段开展档案工作。传统以纸质载体为主的档案工作可用传统的手工管理方式完成，在电子文件大量产生的情况下，档案工作不实现信息化将难以实现对电子文件的归档。因此，档案信息化是企业信息化的必然要求。

## 第二节 数字档案馆定义、特征及发展历程

### 一、数字档案馆定义

对于数字档案馆的定义，不同的文献给出了不同的定义。《数字档案馆建设指南》将数字档案馆定义为，建立在现代信息技术普遍应用基础上，利用数字化手段，以综合档案信息资源为处理核心，对文件和数字档案信息资源进行收集、管理，通过高速宽带通信网络设施相连接和提供利用，实现资源共享的超大规模、分布式数字信息系统。

2017年，国家档案局下发《企业数字档案馆（室）建设指南》，将数字档案馆（室）定义为“企业运用现代信息技术固化档案工作业务流程，对本企业或与其具有资产隶属关系企业的电子档案或其他数字资源进行收集、整理、保存，并通过网络提供档案信息服务和共享利用的集成管理系统平台”。

上述两个定义有一个共同点，就是把数字档案馆定义为信息系统。即数字档案馆本质上既不是一个机构，也不是一座建筑，而是一个由软件、硬件、数据、人、制度规范等组成的人机系统。本书认为，采用信息系统概述的定义符合数字档案馆建设理论，方便企业档案工作者理解。

## 二、数字档案馆基本特征

与传统档案馆相比，数字档案馆具有以下特征。

### 1. 档案资源数字化

这是数字档案馆的本质特征，即管理对象的数字化。档案资源是数字档案馆管理的对象，但与传统档案馆相比，数字档案馆的档案资源在形态上发生了本质的变化，即由纸质为主的传统载体转变为数字档案资源。档案资源形态的转变也正是数字档案馆建设的需求。数字档案馆的数字档案资源主要是通过各类信息系统中形成的电子文件归档和对纸质等传统载体档案进行数字化加工，以数字形式存储各种档案信息。

### 2. 档案管理信息化

信息化是指以现代通信、网络、数据库技术为基础，对所研究对象各要素汇总至数据库，供特定人群生活、工作、学习、辅助决策等和人类息息相关的各种行为相结合的一种技术，使用该技术，可以极大地提高各种行为的效率，为推动人类社会进步提供极大的技术支持。因此，数字档案馆的另一大特征是大量应用现代通信、网络、数据库等技术，极大地提高档案管理效率，尤其是数字档案馆将档案管理业务流程固化在电子档案管理系统中，实现数字档案资源的自动化管理，通过信息技术手段来实现档案的收、管、存、用。

### 3. 档案服务知识化

这里的档案服务知识化包含两层含义，一是将数字档案馆中的数字档案资源库纳入企业知识管理，数字档案资源将作为知识库的一部分，通过数字档案资源丰富企业知识库。二是在数字档案馆建设中，利用知识管理的理念、技术和方法，提升数字档案馆档案服务水平。这里可用的知识管理技术有知识地图技术、知识本体、各种词库技术的运用等。还可以借鉴知识管理的理念，将以用户为中心的理念和主动服务的理念融入数字档案馆建设中。当前，大数据技术如日中天，将它在数字档案馆建设中加以运



用，从而提升档案服务水平。

### 三、数字档案馆发展历程

我国企业档案信息化与全国档案信息化同步发展。我国档案信息化始于 20 世纪 80 年代末，成长于 20 世纪 90 年代，进入 21 世纪以后档案信息化建设迈向稳步成长与快速发展的重要阶段。

#### 1. 启蒙阶段

世界范围内的企业信息化萌生了文档一体化。20 世纪 80 年代中后期，由于市场竞争日益加剧，企业为了提高自身竞争力，纷纷采用现代网络与信息技术开展企业信息化，以帮助降低生产成本，提高产品质量，缩短上市时间，并提供优质服务。20 世纪 90 年代初，办公自动化系统为企业内部流程化管理和高效率运作提供了强有力的技术支持。面对信息化给企业各方面带来的变化，企业档案工作人员也逐步开始了解和接触计算机，并尝试用计算机建立机读档案目录。当时郑州航空工业管理学院开发了一套基于 DBaseIII 的档案管理软件，并在西安飞机工业（集团）公司等企业进行应用。使用该软件可建立档案目录数据库，进行档案数量统计和制作卷内目录、案卷目录等，这个阶段也被称为单机应用阶段。

#### 2. 起步与探索阶段

20 世纪 90 年代，为适应全国各大城市信息化建设和档案事业发展的需要，我国档案信息化工作开始起步。一些省市级档案馆开始考虑档案信息化的规划和建设，并取得了初步成效。《上海市计算机辅助档案管理总体方案》就是这个时期的代表成果之一，一些档案管理机构根据档案管理业务的实际工作需求，研究和开发出单机版的档案管理信息系统（Archives Management Information System）。从 1992 年开始，有一些实力较强的档案馆开始自主开发档案管理信息系统，并取得了初步成果。但是由于当时全国档案行业的信息化基础设施相对薄弱、应用系统功能不完善、档案工作人员知识结构单一等原因，档案信息化规划的一些初步成果也仅停留在计算机辅助管理的局部性和探索性研究阶

段, 档案管理信息系统也局限于单机版对档案目录数据的管理。总体而言, 该时期全国的档案信息化依然处于尝试性的初级应用阶段, 但取得了一定的成果, 1997 年深圳市档案局提出了建立数字档案馆的初步设想, 几年后, 深圳市档案馆成为全国档案工作的示范单位。

### 3. 建设与成长阶段

1995—1999 年, 全国档案工作进入了更大范围的信息化基础设施建设和逐步成长的初级发展阶段。这一阶段的工作重点在于信息化基础建设和应用系统的方案设计。随着档案工作人员的信息化意识逐步增强, 以及档案基础设施与档案资源数据库建设的投入加大, 档案管理软件纷纷出现, 档案信息网也出现在互联网上, 档案目录数据库查询和检索系统在各大城市的档案馆先后建立起来, 利用局域网提高档案信息服务工作也在档案馆内部开展起来; 国家档案局于 1999 年发布了《CAD 电子文件光盘存储、归档与档案管理要求》等相关标准。档案管理系统逐步由单机版发展到客户端/服务器 (C/S) 的网络运行模式。然而, 这一时期的档案信息化工作依然存在诸多问题, 如档案信息化标准、规范相对滞后, 档案管理系统研发人员与档案工作人员之间缺乏沟通, 档案信息系统的应用也处于目录数据集中管理与目录信息局域网内共享的应用阶段, 电子档案全文的收集、保管和提供利用等方面工作尚未涉及。另外, 在馆藏档案数字化与开发利用等方面也未加大投入力度。

但这一阶段的网络化基础设施建设和管理信息系统的使用等工作为“十五”期间全面开展档案信息化建设奠定了良好的硬件基础, 积累了一定经验, 培养了一批档案信息化应用的人才队伍。

### 4. 发展与应用阶段

经过档案信息化的前三个历程, 进入 21 世纪, 我国的电子政务进入实质性运行阶段, 档案信息化建设进入快速发展的应用阶段。2000 年, 国家档案局将加快档案信息化建设列为全国档案事业“十五”规划的一项重要任务, 深圳市率先开始了数字档案馆建设。2002 年国家档案局制定了《全国档案信息化建设实施纲要》, 要求对档案管理实行网络化和自动化, 要求各档案馆建立自己的档案信息网, 向外公开馆藏的一些数据。2003 年青岛市数字档案馆正式启用, 成为我国首家投入使用的数字档案馆。当时, 我国的档案信息化进入了一个快速

发展时期，电子档案管理系统逐步建立起来且功能相对完善，并进一步发展到浏览器/服务器（简称 B/S）模式，国家档案局制定并发布了《电子文件归档与管理规范》、《档案管理软件功能要求暂行规定》。企业信息化产生了大量的电子文件，有些电子文件很难以纸质形式进行保存，如果以电子形式归档，必然形成电子形式的档案，档案保管机构面临如何接收和保管电子文件以及如何提高利用效率的挑战，电子文件的归档管理开始得到重视。

## 5. 深化应用阶段

随着电子文件产生的数量越来越多，数字档案馆建设开始由单元技术突破，向集成化功能方向发展，管理重点由传统载体档案转向电子档案。数字档案馆的功能进一步拓展，对提升档案管理水平起到了重要作用。在此期间，国家档案局发布了《数字档案馆建设指南》、《数字档案室建设指南》、《企业数字档案馆（室）建设指南》等指导性文件，数字档案馆建设成为档案信息化建设的方向和主要内容。

# 四、数字档案馆在我国企业的应用

2004 年，我国第一个企业数字档案馆——江苏省电力公司数字档案馆作为国家档案局“十五”科研项目在南京通过验收，该项目在馆藏数字资源的综合利用、多格式安全浏览、实现数据安全等方面均处于国内领先和国际先进水平。国家电网以此为基础开始建设覆盖所有供电局和电力公司的数字档案馆网络。

2006 年，当时的中国航空工业第一集团公司提出，要在 2010 年前建成全集团的数字档案馆，同年制订了指导集团各成员单位开展数字档案馆建设的工作方案，并于 2007 年推出了第一个企业数字档案馆建设评价体系，通过评分制考核数字档案馆建设水平，根据得分多少，确定航空工业一级数字档案馆、二级数字档案馆和三级数字档案馆。2010 年，航空工业档案馆分别对 5 家单位的数字档案馆进行了评定，并确定相应等级的数字档案馆。

2010 年之后，电子文件管理在企业得到了前所未有的重视。不少企业已开始探索档案的知识化管理，韶钢集团、中远天津散货等企业的档案知识化管理取得了较好的效果。企业数字档案馆系统软件产业竞争进入白热化，软件厂商

达到上百家，每年产值达十几亿元。大型企业集团档案信息化开始由分散走向集中，电子档案管理系统的部署由各自为战变为集中部署。档案信息化建设的知识化趋势明显，不少企业提出并着手实施企业档案信息的知识化管理，通过管理与技术的结合，实现档案信息的知识化管理。

随着信息化技术在企业更加深入广泛的运用，电子商务“十三五”规划、互联网+、“中国制造 2025”等国家战略的实施，企业产生的电子文件数量进一步增加，企业对数字档案馆的升级换代需求开始变得更加迫切。为此，国家档案局在制定《全国档案事业发展“十三五”规划纲要》时提出，要“全面推进档案资源存量数字化、增量电子化、利用网络化”，要“开展企业示范数字档案馆建设，建成一批具有国际先进水平的企业数字档案馆”。根据这一规划，国家档案局于“十三五”期间大力推进企业传统载体档案数字化工作，力争在 2020 年实现企业传统载体档案数字化率达到 70% 以上。同时，在企业电子文件归档与电子档案管理试点的基础上，着力解决业务类电子文件归档难题，提升档案信息开发利用水平。2017 年国家档案局制定了《企业数字档案馆（室）建设指南》，对企业档案数字档案馆（室）建设的目标、设施设备要求、电子档案管理系统功能、传统载体档案数字化、数字档案馆建设方法步骤等做出了规定。2018 年初，经过企业申报和专家评审，国家档案局确定了 35 家企业数字档案馆（室）建设试点单位。目前，试点工作正在稳步推进之中。

### 第三节 企业数字档案馆总体架构

#### 一、研究企业数字档案馆总体架构的意义

第一，与机关档案室、综合档案馆相比，企业建设的数字档案馆在业务流程和管理对象上有其特殊性，其业务流程与综合型档案馆有较大差异。特别是大中型企业档案馆，管理对象多种多样，有文书档案、科技档案，而科技档案又有产品档案、科研档案、建设项目档案等，种类多达 20 余种。从业务流程来看，企业档案馆不仅要管理档案，还要对科技文件材料进行管理。管理对象和业务流程的差异，使其在数字档案馆建设时的总体框架和建设内容与综合档案馆、机关档案室有较大区别，大多数企业尚不能十分准确把握企业数字

档案馆的总体框架模型和建设内容，更无法确定建设目标。

第二，企业数字档案馆总体框架决定数字档案馆建设的内容，对数字档案馆建设的质量起着决定性作用。尽管企业数字档案馆建设工作内容繁多，但从本质上来说是一个信息系统的设计与实施，任务信息系统的设计与实施都是从总体框架自顶向下逐层设计的，这是任务信息系统通常采用的生命周期法。信息系统的总体框架决定了其所属模块的结构和细节，同时也决定了系统建设的内容，这些都是决定数字档案馆质量的关键因素。因此，企业建设数字档案馆必须先设计总体架构，再由总体架构派生出建设细节和建设内容。

第三，一个统一的数字档案馆总体架构模型有利于数字档案馆的档案信息共享与系统集成。数字档案馆在本质上是一个任务信息系统，其总体架构决定了其详细的体系结构，而体系结构又决定不同企业数字档案馆的信息共享与系统集成可否实现。体系结构相同的数字档案馆可方便地实现信息共享和系统集成，相反则难。在一个统一的数字档案馆总体架构模型指导下建立起来的数字档案馆，其体系结构是基本相同的，易于实现数字档案馆的信息共享和系统集成。数字档案馆之间的信息共享与系统集成程度是决定数字档案馆作用能否充分发挥的重要因素，存在信息孤岛的数字档案馆是难以充分发挥作用的。特别是一些国有大中型企业、集团公司，有下属子公司，如果同一集团或行业所建的数字档案馆由于总体架构不同，而使体系结构差异较大，造成数据交换和信息共享不畅，更不能与集团公司进行集成，其作用就会大打折扣。

## 二、企业数字档案馆总体架构的层次模型及其功能

通过以上论述可知，数字档案馆在本质上是一个信息系统，且信息系统是有架构的，因此，数字档案馆总体架构可以简单定义为，数字档案馆各部件的构成框架。由于对部件的不同理解可构成不同的架构方式，如概念架构、功能架构、硬件架构等。而数字档案馆是一个由硬件、软件、数据库和人员等组成的综合体，用功能架构和硬件架构都无法充分表述它的组成部分，因此，这里只讨论概念架构。运用概念架构，对数字档案馆各组成部分进行从低到高的分层抽象，形成了数字档案馆的层次架构模型，即数字档案馆组成部分的层次结

构关系。

与传统的档案馆相比，数字档案馆在结构上有着本质的不同。按照层次来划分，数字档案馆架构从下到上依次为基础层、数据层、管理应用层、运行机制层和贯穿全过程的标准规范体系、安全保密体系六个部分构成，具体见图 1-1。

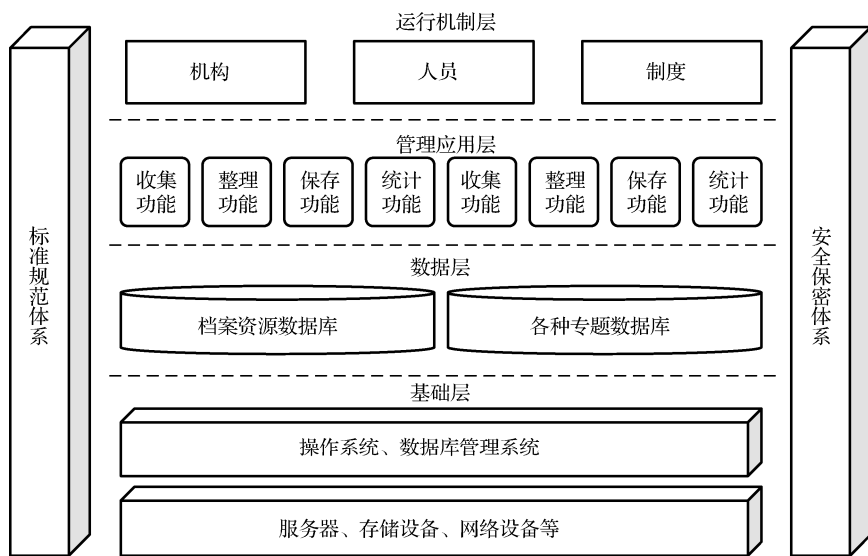


图 1-1 企业数字档案馆总体架构图

## 1. 基础层

基础层是数字档案馆的最底层，主要包括硬件设备和系统平台。硬件有网络、服务器及相应的存储设备。与硬件处于同一层的还有系统平台，包括操作系统和数据库管理系统，其中数据库管理系统又是以操作系统为平台，并处于操作系统之上。硬件和系统平台是整个数字档案馆最基础的部分。

## 2. 数据层

数据层是数字档案馆管理的对象，按照数据来源可分为电子档案、其他数字资源和专题档案数据三部分。

### 3. 管理应用层

管理应用层是数字档案馆的核心，即电子档案管理系统。

### 4. 运行机制层

运行机制层指的是数字档案馆运行所依赖的组织机构、管理制度、人员等，它是数字档案馆发挥效用的重要组成部分，它位于数字档案馆的最顶层，对数字档案馆的运行具有约束作用。

### 5. 标准规范体系

标准、规范是实施数字档案馆项目的重要基础，也是数字档案馆日常运行必不可少的组成部分。面对数字资源形式的多样性，社会对数字资源共享要求的广泛性以及相关技术、设备不断发展、变化的动态性特点，结合建设内容，根据国际、国内相关标准，逐步推出数字档案馆建设标准规范，并形成标准规范体系。按照标准规范的内容可划分为管理性标准规范、业务性标准规范、技术性标准规范三类。

### 6. 安全保密体系

数字档案馆是运用信息技术对档案信息进行管理，因此，数字档案馆的建设和运行必须有完善的安全保密体系作为保障。

数字档案馆的安全包括网络安全、系统安全、信息安全、物理安全等方面，要通过可靠的技术措施和完善的管理制度来确保各方面的安全。

## 三、数字档案馆总体架构层次模型的启示

数字档案馆总体架构的层次模型，不单是一个数字档案馆抽象的概念模型，从这个模型中可以反映出诸多重要信息。

第一，总体架构的层次模型可延伸出数字档案馆的建设内容。总体架构层次模型是一个由数字档案馆各组成部分按层次抽象出来的模型，这个模型能直观地映射出数字档案馆的建设内容。与该模型各层相对应的建设内容见表 1-1。

表 1-1 各层相对应的建设内容

层次	建设内容
基础层	硬件购置、网络搭建、操作系统购置与安装、数据库管理系统购置与安装
数据层	传统载体档案数字化、电子文件归档、专题数据库建设
管理应用层	电子档案管理系统开发、实施等
运行机制层	数字档案馆运行所必需的制度、组织机构、人才队伍建设
标准规范体系	数字档案馆建设与运行管理必需的标准规范建设
安全保密体系	数字档案馆运行所必需的安全与保密措施建设

第二，总体架构的层次模型显示了数字档案馆各组成部分在数字档案馆中的地位与作用，有利于建设单位做出正确的规划与设计。数字档案馆各组成部分是一个自底向上依次分步的层次结构，其各部分在数字档案馆总体架构中所表现出来的作用各不相同。处于底层的组成部分是基础，中间层是核心，上层是重点，贯穿各个层次的是保障。正是由于各部分所处的地位不同，使建设单位在制定数字档案馆建设规划时，可遵照“一步规划，分步实施”的原则，区分轻重缓急，有计划有步骤地安排各部分的建设，从而避免一次性投资过大，有利于降低数字档案馆建设的风险。

第三，总体架构层次模型为构建数字档案馆建设的评价指标体系奠定了基础。有关数字档案馆建设的评价问题尚未引起足够重视，但随着这项工作的深入开展，数字档案馆建设水平的高低将被提上议事日程。要对数字档案馆建设水平进行评价，就必须构建数字档案馆建设评价指标体系。如果缺乏数字档案馆的总体架构层次模型，其建设内容也无法确定，进而直接导致建设水平评价指标体系难以构建。反之，如果通过分析数字档案馆的总体架构层次模型，明确数字档案馆各组成部分及其相互关系，在此基础上再开展评价指标体系研究就容易多了。

## 四、企业数字档案馆建设内容

建设内容是企业数字档案馆建设的重点，它决定了企业数字档案馆建设的任务、投入、过程等。根据前一节企业数字档案馆的总体框架层次模型，企业数字档案馆的建设内容可分为基础设施建设、电子档案管理系统开发实施、数字档案资源建设、运行机制建设、标准规范体系建设和安全保密体系建设六个



部分，具体见图 1-2。

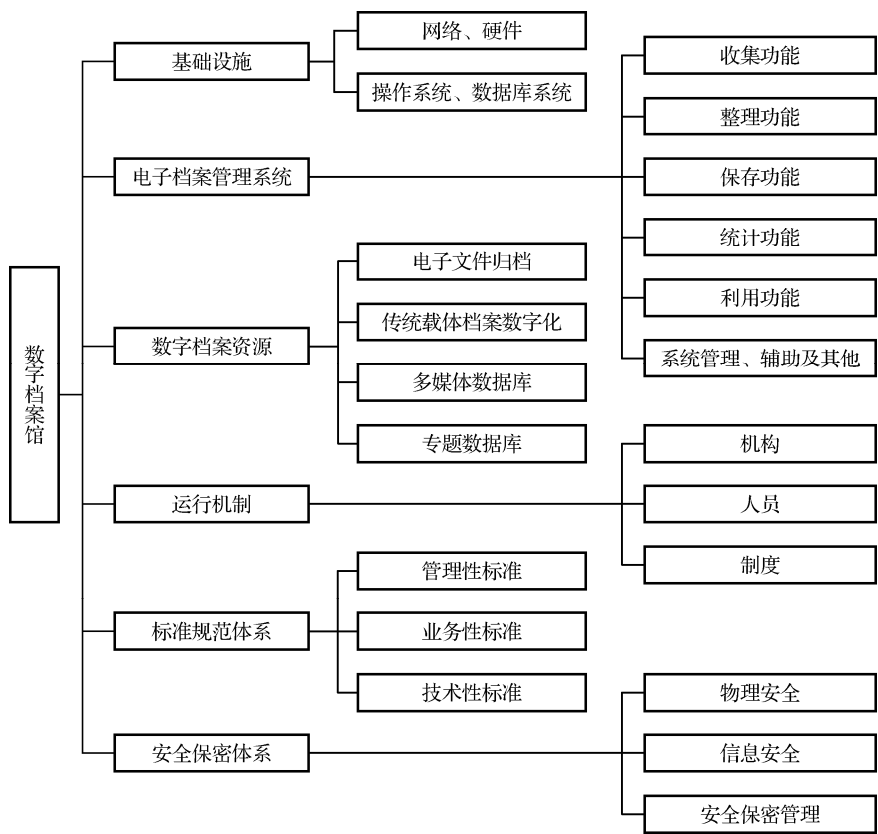


图 1-2 企业数字档案馆建设内容图

1，基础设施建设

企业数字档案馆的基础设施建设包括场地环境建设、网络设施建设、服务器、数据存储备份设备和终端硬件设备配置、系统软件购置等。

(1) 场地环境建设

场地环境建设是整个数字档案馆建设的基础，应保证机房、数字化中心、电子阅览室等配套场地、温度、湿度、光照强度等环境因素达到硬件所要求，配备门控系统、监控报警系统和区域保护设施等，配备磁带备份系统、光盘记录系统、断电保护 UPS 系统等外围辅助设备，确保环境安全。

### （2）网络设施建设

数字档案馆网络是企业整个网络结构的有机组成部分，应统筹规划、设计和建设。网络设施的配置要充分考虑电子文件采集、归档和数字档案安全管理、存储数据容量、数据操作强度、网络类型、网络访问量以及档案信息网络利用服务、数据迁移等工作要求。网络性能应能适应图像、音频、视频等各类数据的传输、利用要求。

### （3）服务器、数据存储备份设备和终端硬件设备配置

服务器配置应能满足数字档案馆（室）档案管理系统运行、数据备份服务等的需求；应配备稳定与适当冗余的在线、近线与离线数据存储设备，配置科学、可行的存储策略，确保档案资源长期保存与有效利用；应配备档案专用的计算机终端、扫描仪、数码照相机、打印机和刻录机、移动存储介质等设备。

### （4）系统平台购置

应结合企业信息化的现状，统筹协调利用企业现已运行的基础系统平台，购置运行电子档案管理系统所需要的操作系统、数据库管理系统等系统平台。

## 2. 电子档案管理系统建设

电子档案管理系统建设是企业数字档案馆建设的核心内容，主要包括开发功能完善、安全可靠、性能良好、方便易用的电子档案管理系统，实现档案的收集、整理、保管、统计、利用等各业务环节的自动化、网络化。电子档案管理系统应与企业办公自动化、财务会计、人力资源等各业务系统做好衔接，具备各类电子文件归档与电子档案管理的功能。

## 3. 数字档案资源建设

数字档案资源建设是企业数字档案馆建设的重要任务，一般包括传统载体档案数字化、电子文件归档和专题数据库建设。

### （1）传统载体档案数字化

一定数量的档案数据是保证数字档案馆发挥效用的基础，存量档案数字化率宜达到可数字化范围的70%以上，传统载体档案数字化的内容和范围可结合企业实际情况确定。

### （2）电子文件归档

电子文件归档范围应包括企业各项业务活动中形成的各种结构化、非结构

化数据和以独立文档形式存储的具有保存价值的信息记录，包括办公自动化系统、产品或业务系统、财务会计管理信息系统、人力资源管理信息系统、门户网站、微博、微信、公务邮件系统和本企业其他职能活动业务系统，以及从外部接收的电子文件。企业有些信息系统有可能通过租用基础设施或以云计算服务的形式存在，所形成的电子文件不一定存在本企业的服务器中，但其产生的电子文件也应纳入本企业文件材料归档范围予以归档。企业已实施的支撑主营业务的信息系统均应具有归档功能，导出的归档电子文件存储格式、元数据等均应符合电子文件归档和电子档案管理的有关要求。由于国内电子文件归档技术尚在逐步完善，数字档案馆建设对电子文件归档的初步要求是完成办公自动化系统（OA）形成的电子文件归档并实现核心业务系统电子文件归档。

### （3）专题数据库建设

专题数据库建设是在现有数字档案资源的基础上，通过有组织的分析、筛选、整合，把某一特定专题的档案集中、有序、系统地组织在一起，以充分发挥数字档案资源价值、实现数字档案资源的高效利用和共享。

## 4. 制度规范建设

制度规范建设是企业数字档案馆建设的一项重要基础性工作，包括管理制度建设和标准规范建设。

### （1）管理制度建设

企业数字档案馆建设应制定数字档案馆人员岗位设置与职责制度，系统运行维护管理制度，服务器机房、档案数字化加工场所管理制度，数字档案馆安全与保密管理制度等。

### （2）标准规范建设

企业数字档案馆建设应参考国家相关标准规范并结合本企业实际，形成本企业基于数字档案馆的电子档案收、存、管、用等方面的标准规范体系，包括但不限于以下内容。

① 企业传统载体档案数字化加工标准规范，包括全文数字化加工和目录数字化著录细则等，全文数字化可参照《纸质档案数字化规范》（DA/T31）制定，缩微胶片数字化参照《缩微胶片数字化技术规则》（DA/T43）制定。企业在委托中介机构开展档案数字化外包服务时，应严格遵守档案信息保密安全的有关规定。

② 电子文件归档标准规范，主要规范本企业已实施的各类信息系统产生的

电子文件归档业务，内容包括各类信息系统产生电子文件的归档管理责任、归档范围、归档流程及开发利用要求等。

③ 企业档案分类规范，根据企业档案自身特点制订企业档案的分类方案，以供数字档案馆建设过程中对档案信息存储进行组织时参照。

④ 企业各类电子档案管理及业务流程规范，包括管理类、产品/业务、科研、基本建设、设备仪器、会计、职工管理等各类档案信息在电子档案管理系统中采集、整理、存储、保管和利用等的规范。

⑤ 企业档案利用标准规范，包括企业档案利用赋权原则、权限设置与更改流程，档案利用审批流程，计算机辅助编研规范等。

⑥ 企业数字档案馆安全与保密管理系列规范，具体参见第九章。

## 5. 安全保密体系建设

数字档案馆（室）建设必须加强安全管理，遵守相关保密规定，安全设备配置齐全，确保物理安全；严格控制信息资源利用范围，制定严密的信息发布审核制度，确保信息安全。数字档案馆系统应具备数据的防灾备份功能，严格控制数据访问权限，制定完善的数据库管理规范，确保数据安全。

管理涉密档案的电子档案管理系统，应按照《涉及国家秘密的信息系统分级保护技术要求》（BMB17）进行安全设计和建设，并通过有关部门测评后方可投入使用。

管理非涉密档案的电子档案管理系统应根据系统重要程度及实际安全需求，按照《计算机信息系统安全等级保护划分准则》（GB/T17859）确定档案管理系统应达到的安全等级，并依照《信息安全技术信息系统安全等级保护基本要求》（GB/T22239）、《信息安全技术信息系统安全等级保护定级指南》（GB/T22240）、《信息安全技术信息系统安全等级保护实施指南》（GB/T25058）开展安全建设和登记管理。企业使用的电子档案管理系统一般为非公共系统，安全等级一般为二级或三级。管理非涉密档案的电子管理系统应与互联网逻辑隔离。

## 6. 经费与人才保障

### （1）经费保障

经费主要包括计算机硬件采购费，软件开发、升级、购置、接口开发、培训、评测费，运行维护费，传统载体档案数字化费，项目实施费，人员培训费

等。企业数字档案馆建设应按照整体规划，制定相对应的经费计划，确保管理经费、实施经费及软硬件设备购置费的落实，为企业数字档案馆建设工作提供必要的保障。

## （2）人才保障

应结合本企业业务实际积极开展档案部门人员业务和信息化能力的培养及相关专业人才的引进，以满足数字档案馆建设和运行的需求。在人才引进方面，应注重档案和信息化专业相结合，建成一支既懂档案业务又具有一定信息化技能的复合型人才队伍。

各部分详细的建设内容请参考后续章节。

## 第四节 企业数字档案馆建设步骤

根据企业数字档案馆的定义，它的六个组成部分构成一个概念模型，如何将概念模型转化为实际的物理模型，这个过程就是建设过程。在具体实践中，企业数字档案馆建设按照建设内容和先后顺序一般按照启动（规划与立项）、实施、验收、维护等步骤分阶段进行，具体见图 1-3。

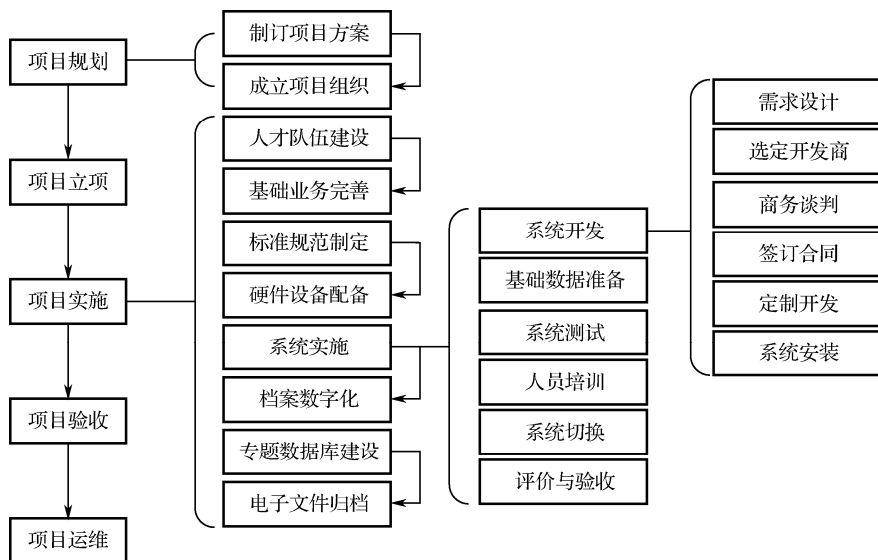


图 1-3 企业数字档案馆建设过程图

## 一、项目规划

企业数字档案馆建设是一项系统性工程，涉及企业的方方面面，必须做好规划。通过规划，分析企业信息化现状、档案工作现状、档案信息化现状，继承企业发展目标和档案工作发展目标，分析企业建设数字档案馆的内外部环境，确立标杆，再确定企业档案数字档案馆建设目标、任务、实现路线和实施效果的过程。因此，数字档案馆规划的过程也是企业进行数字档案馆建设可行性论证分析的过程。

## 二、项目立项

规划方案是编制者智慧的结晶，但它仅代表编制者的观点，受制定者自身学识、观念、认知、经验等因素限制，规划方案并不一定可行。为了使规划方案更加科学合理，应聘请单位内外专家对规划方案的可行性、科学性进行评审论证，从而使它在作为后续工作依据时更加科学。

经过规划设计的项目方案仅仅是一张蓝图，要把蓝图转换为现实，还必须要有经费作保障。企业数字档案馆建设要根据规划中确定的经费需求，通过立项，将经费纳入预算。一般来说，可将数字档案馆建设经费纳入两种预算，一是纳入企业信息化预算，二是纳入企业档案工作的固定资产投资预算。经费来源可以是国拨，也可以是自筹，或者国拨与自筹相结合。

## 三、项目实施

项目实施就是把规划方案中确定的蓝图转化为现实的过程，这是数字档案馆建设的关键。它由前后相继的一系列过程组成。项目实施是数字档案馆建设过程中用时较长、花费较高的，需要投入较多人力物力，企业必须精心组织好本阶段工作。

## 四、项目验收

经过项目实施工作，数字档案馆建设基本实现了从蓝图到物理的转化。但这个转化是否成功，是否达到预期目标，需要通过验收来确认。项目验收是数字档案馆建设项目的里程碑节点，即对项目实施过程所有工作的总结评价。要通过验收，总结实施工作的得与失，对实施工作是否达到规划方案所确定的目标和任务进行评判，为后续工作提供依据。

## 五、项目维护

数字档案馆建设是基于建设期间企业的内外部环境、资源等约束条件来实施的。但随着企业内外部环境的变化，数字档案馆的建设需求也会发生改变，如功能模块的增删或修改、应用范围的扩大或缩小、软件硬件性能的提升等。数字档案馆要应对这些变化，必须不断地进行升级和维护。同时，数字档案馆作为一个信息系统，在运行过程中也同其他信息系统一样需要进行管理，要有专门的人员对出现的问题进行处理，对故障进行排除，还要有专门的经费为功能模块的更改、系统性能的提升等提供保障等，只有这样，才能保证数字档案馆系统的稳定运行和正常使用。

# 第五节 企业建设数字档案馆的效益

## 一、顺应档案工作发展的必然要求

21 世纪，信息技术的飞速发展推动生产力的发展超过了以往的任何一个时期，信息技术已经成为当代社会最活跃的生产力。党的十六大会议就明确指出：“信息化是我国加快实现工业化和现代化的必然选择”。以信息化带动工业化，以工业化促进信息化已经成为我国现代化建设的必由之路。

信息技术的飞速发展，为企业的现代化、市场化提供了重要条件，为实现

企业发展战略、技术创新战略和市场营销战略提供了有效保证。信息时代企业盈利的法则是把信息变成知识，把知识变成决策，把决策变成利润。为此，信息化水平已成为现代企业竞争力和生命力的重要标志。信息技术的飞速发展和广泛应用，对国民经济和社会发展产生了深刻影响，同时对档案工作提出了更高的要求。信息化为档案事业发展带来了前所未有的挑战与机遇。档案工作作为企业基础管理的一个重要环节，应用信息技术全面提升档案工作水平是档案事业适应时代和社会发展的必然选择，是加速档案管理现代化的客观要求。

企业档案信息化是企业信息化的重要组成部分，没有档案信息化的企业信息化是不完整不彻底的信息化。企业档案信息化带来的好处不仅惠及企业档案工作本身，更重要的是它能保证企业信息化过程中出现的数据安全不出现问题，解决企业信息化的后顾之忧。大力加强企业档案信息化工作是档案生产力发展的必然要求，是档案工作发展的必然选择。

## 二、助力企业信息化建设

自 20 世纪 80 年代中国企业开始大量应用信息技术至今，形成了大量的电子数据。国际数据公司（IDC）的研究结果表明，2008 年全球产生的数据量为 0.49ZB，2009 年增长至 0.8ZB，2010 年再增长至 1.2ZB，2011 年则高达 1.82ZB，相当于全球每人产生 200GB 以上的数据。截至 2012 年，人类生产印刷材料的数据量约为 200PB，人类说出语言的数据量约为 5EB。据 IBM 的研究称，整个人类文明所获得的全部数据中，有 90% 是过去两年内产生的。而到了 2020 年，全世界所产生的数据规模将达到今天的 44 倍。这么多的数据，我国企业却一直保存在原始系统中，不仅增加系统运行负担，还给系统升级带来了诸多困难。有的企业由于系统数据量太大，致使不敢扩大系统应用范围，有的企业因怕系统升级后大量数据无法迁移，致使企业不敢实施系统升级，而有些企业强行升级，致使数据在迁移过程中，大量数据丢失或损坏。有些企业为了减轻系统运行负荷，强行将数据以备份形式移出，给这些数据长期保存带来风险。有的企业由于没有实施数字档案馆，使电子文件无法归档，会计档案电子化无法实施，电子发票无法接收，企业电子商务和业务流程无法得到优化。

数字档案馆建设将电子文件的归档管理作为重点内容，应用电子数据归档技术，使企业现有数据得到长期可靠安全的存储和便利的利用，企业数据风险



得以化解,促进系统应用范围根据企业需要扩充,降低企业信息系统运行负荷,促进了企业信息化建设。

某企业实施数字档案馆时,将会计电子文件归档作为重点内容,实现了凭证、账本、报表的电子化管理,特别是实现了电子发票的电子化管理,不再开具和接收纸质发票,仅此一项,每年为企业节约管理成本近亿元。

### 三、实现档案事业的平衡协调发展

中国共产党第十九次全国代表大会提出,我国社会主要矛盾由“人民日益增长的物质文化需要同落后的社会生产之间的矛盾”转化为“人民日益增长的美好生活需要和不平衡不充分的发展之间的矛盾”。未来一段时间工作重点就是要解决发展不平衡、不充分的问题。具体到企业档案工作,随着经济发展和企业改革的深化,方方面面对档案工作提出了迫切需求。横向比较来看,全国企业档案工作发展也极不平衡。要解决企业档案工作中这些问题,一条重要措施就是建设数字档案馆和数字档案室,实现企业档案工作的数字化转型。因为,数字档案馆和数字档案室具有提高档案管理效率,提升档案利用水平等作用,通过它的建设,能有效地解决企业档案工作发展不充分、不平衡的问题。通过数字档案馆和数字档案室建设,能使档案信息化与企业信息化发展水平相适应,与企业档案信息需求相适应,最终能使各单位档案信息化水平与国家档案信息化整体水平相适应。这是落实“十三五”规划,提升档案信息化总体水平的需要。2016年,国家档案局制定的《全国档案事业发展“十三五”规划纲要》,提出“十三五”期间“要持续推进数字档案馆建设”,“开展企业示范数字档案馆建设,建成一批具有国际先进水平的企业数字档案馆”。为了落实这项规划,2017年,国家档案局印发了《企业数字档案馆(室)建设指南》,加强企业数字档案馆和数字档案室建设,并持续推进此项工作。通过《企业数字档案馆(室)建设指南》可知,如今的企业数字档案馆和数字档案室,远不是前几年的景象,着眼于建设“具有国际先进水平的企业数字档案馆”,是升级换代了的数字档案馆和数字档案室,是数字档案馆的3.0或4.0版。推进此项工作就是要让企业档案信息化升级换代,跟上信息化时代发展的步伐。国家档案局希望通过推进此项工作,真正建成一批具有国际先进水平的企业数字档案馆和数字档案室,引领我国档案信息化建设走向高端,走在世界前列。

## 四、解决当前企业档案工作面临的困难与问题

尽管新中国成立以来，我国的档案事业取得了长足的发展。但在当前市场经济的大潮下，档案工作也面临着前所未有的挑战。企业档案工作边缘化的趋势仍旧存在，档案机构撤并现象在很多企业发生。仅在军工集团，企业档案部门由独立机构合并为非独立机构的情况每年都有发生，而对档案机构进行重新独立设置的企业少之又少。档案工作与企业财务、人力资源、研发、市场销售等工作相比地位有进一步下降的趋势。这种趋势的发生有档案工作本身的后端性、创造价值的滞后性等原因，同时也有企业档案工作应用新技术滞后的原因。一个最好的对比便是图书情报工作，长期以来，图书情报工作在企业也存在被边缘化现象，不少企业图书情报人员纷纷换岗。但在信息技术应用的大潮来临之时，图书情报工作较好地抓住了机遇，大量应用信息技术，并较早提出了数字图书馆的概念，应用信息技术大幅度提高了信息的检索效率，使之服务企业主营业务的效率不断提高，工作重新得到企业的重视，尤其在高新技术企业，图书情报得到了与技术研发并驾齐驱的地位。因此，面对当前的困难与挑战，档案工作也必须积极应用信息技术，实现全流程全方位的信息化，通过信息化提高档案工作效率，加强档案信息的开发利用工作，提高档案工作的技术含量，丰富档案工作的内容，并最终成为企业创新效益的新生力量。

## 五、实现企业档案提质增效

数字档案馆将档案管理业务流程固化在电子档案管理系统中，实现数字档案资源的自动化管理，大幅度提高档案管理效率，主要表现在以下方面。

### 1. 数字档案馆明显提升企业档案管理效率

数字档案馆将档案管理业务流程固化在电子档案管理系统中，实现文件的收集、鉴定、整理和数字档案资源保管、利用等的自动化，管理效率明显提升；有的数字档案馆应用物联网技术还实现了传统载体档案的自动传输、自动盘点，调卷效率和盘点效率明显提高。经过统计，实施数字档案馆之后，档案管理效

率是原先的 4 倍，需要 4 人管理的档案只需 1 人就能胜任。

## 2. 数字档案馆明显提升企业档案管理质量

### (1) 数字档案馆提升文件归档鉴定质量

数字档案馆将保管期限表嵌入电子档案管理系统，通过多种匹配技术，实现保管期限划分的自动化，避免了由于不同档案管理员在技能上的差别形成的保管期限划定的差异和人为干扰，使档案保管期限划定一致性大为增强，归档时鉴定质量明显提高。

### (2) 数字档案馆提升文件整理质量

数字档案馆将文件整理规则嵌入电子档案管理系统，大幅度降低档案管理人员对文件整理规则理解的偏差，使文件整理质量大为提高。应用电子档案管理系统，实现元数据自动采集，可大幅度降低元数据错误率。

### (3) 数字档案馆提升档案保管质量

在档案保管方面，数字档案馆实现档案出入库、人员出入库的自动化登记，档案安全得到有效保证；数字档案馆实现温湿度自动记录和调控，使档案保管环境得到持续维护，有利于档案保管。

## 3. 数字档案馆使档案信息在企业经营管理活动中的支撑效应得到放大

数字档案馆建成大幅度提高档案信息检索速度，提高档案信息查询查全准确率，使员工检索档案更加方便，员工利用档案的积极性进一步提高，使档案信息在企业经营管理活动中的支撑效应得到放大。

如某企业在实施数字档案馆前，利用一份档案需经过目录查找、手工审批、复印等 5 至 6 个环节，从申请到取得档案要 4 个小时，档案利用极为不便，每年档案利用只有区区 50 多人件次。但在建成数字档案馆之后，员工通过网络利用档案，档案信息检索、申请、审批等均在网上完成，利用时间只有不到 1 分钟，有的档案信息获取时间甚至下降至秒级，从而大幅度降低企业档案信息获取成本。由于员工利用档案更加方便，员工利用积极性明显提高，工作中遇到难题首先找档案，档案利用量增长了 10 倍，电子档案管理系统成为员工离不开的系统，档案为企业经营管理提供了大量的支撑和服务，得到了员工的好评，也为企业创造了可观的经济效益。假设一件档案产生的效益为  $a$ ，以前一年利用 50 件档案产生的总效益为  $50a$ ；现在档案利用数量为 500 件，则其产生的效

益应为 500a。数字档案馆给企业带来的这种效益实质上是网络培增效应。

#### 4. 现实案例进一步证明数字档案馆的效益

自 20 世纪 80 年代初企业应用信息技术开始,企业档案工作也尝试应用信息技术,提升档案工作水平。经过几十年的发展,不少企业的档案信息化工作取得了喜人的成绩,为企业的经营与发展提供了有力的支持。中国电信集团、国家电网公司、中远天津散货、本钢集团等企业在实施档案信息化之后,档案工作效率得到了大幅度提升,档案工作较好地支持着主业的发展,档案工作水平也有了大幅度提升,档案部门地位有了较大的提高,档案人员得到了成长。

## 第六节 数字档案馆与档案信息化阶段理论

### 一、诺兰阶段模型及其指导作用

#### 1. 诺兰阶段模型

信息化技术进入一个组织的应用,一般要经历从初级到高级的成长过程,诺兰(Nolan)于 1973 年首次提出信息系统发展的阶段理论,被称为诺兰模型;1980 年,诺兰进一步完善模型,将信息系统的成长过程划分为六个不同的阶段。

(1) 初装:组织购置第一台计算机并初步开发管理应用程序。计算机的作用被初步认识到,个别人具有了初步使用计算机的能力。一般而言,“初装”阶段大多发生在单位财务、人事等数据处理量大的部门。

(2) 蔓延:应用初见成效,信息系统从少数部门扩散到多数部门,并开发了大量的应用系统,使组织的事务处理效率更高。同时出现了数据冗余性、不一致性以及难以共享等有待解决的问题,只有一部分计算机的应用收到了实际效益。

(3) 控制:计算机数量超出控制,预算每年以 30%~40%或更高的比例增长,投资的回报不理想。应用项目不断积累,要求加强组织协调,出现了由组织领导和职能部门负责人参加的领导小组,对整个组织的系统建设进行统筹规

划，特别是利用数据库技术解决数据共享问题。该阶段是实现从以计算机管理为主到以数据管理为主转换的关键。

（4）集成：在控制的基础上，对子系统硬件进行重新连接，建立集中式的数据库及能够充分利用和管理各种信息的系统。重新装备大量设备，预算费用又一次迅速增长。

（5）数据管理：集成阶段之后，将会进入数据管理阶段。

（6）成熟：成熟的信息系统可以满足组织中各管理层次（高层、中层、基层）的要求，从而真正实现信息资源的管理。

## 2. 诺兰阶段模型的作用

诺兰阶段模型总结了管理信息系统的经验和规律，其基本思想对于管理信息系统的建设具有指导意义，一般认为模型中的各个阶段都是不能跳越的，无论是确定开发管理信息系统的策略，还是制定管理信息系统的规划，都应该首先明确组织当前处于哪一个成长阶段，进而根据该阶段特征来指导管理信息系统的建设。

## 二、档案信息化阶段的划分

对于档案信息化发展阶段的划分，作者检索了有关文献，国内尚无专门的研究，本书借鉴诺兰的阶段模型设计出档案信息化发展阶段模型。

由于档案信息化是在企业信息化发展到一定阶段之后才发展起来的，因此，档案信息化发展阶段与企业信息化的发展阶段并不完全相同，而要通过档案机构或企业档案信息化过程的观察及调研并结合诺兰的阶段模型来进行划分。档案信息化也有一个初装阶段，但由于档案信息化是在企业信息化的基础上发展起来的，档案信息化的初装阶段比企业信息化的初装阶段信息化程度更高。档案信息化也有蔓延阶段，即由个别应用向多方面应用发展。档案信息化由于是一个专业领导的信息化，基本不存在集成应用阶段，但存在一个功能集成的阶段，这个阶段的典型特征就是档案工作的全过程信息化，并重视数据管理，因此档案信息化的第三阶段可视为诺兰模型的第三、四阶段的合并。档案信息化接下来一个阶段可借用诺兰阶段模型，称之为成熟阶段，这个阶段就是当前正在研究的数字档案馆阶段，即实现了档案工作过程和管理对象的全数字化。

因此,可将档案信息化划分为五个阶段,从低到高依次为单机应用阶段、业务工作计算机化阶段、工作协同阶段、数字档案馆阶段和高级阶段,各阶段特征如下。

(1) 单机应用阶段。属于档案信息化引入期,此时档案部门刚刚接触信息化,零星地配备一台或若干台计算机,尚未购置专用的档案管理软件,但可通过电子表格等工具对目录进行初级的管理,包括表格输出、目录查询等。

(2) 业务工作计算机化阶段。由于档案数量的增加,或是受益于单机应用,档案部门开始购置软件,将一些手工的工作搬到计算机上,通过计算机、档案管理软件进行档案目录的输入、输出、查询等工作,档案工作效率大大提高;意识到档案全文数字化的效用,着手开展档案全文数字化工作,开始重视数据管理;技术力量较强的单位还将档案信息对内发布,供员工查询。

(3) 工作协同阶段。得益于业务工作计算机化的效益,档案部门着手档案工作所有环节的信息化,实现档案工作的协同,购置网络版的档案管理软件;为了实现所有业务的计算机化,已开始重视管理制度和标准规范的建设,注重业务流程的重组,并贯彻文档一体化和前端控制的档案工作理念;档案目录数据、全文数据基本实现数字化,开始重视电子文件的管理。

(4) 数字档案馆阶段。档案部门开始全面建设企业数字档案馆,档案信息数字化率比较高,档案工作环节基本在信息系统中完成,档案信息化管理重点由存量档案的数字化转变为增量的电子档案管理。这个阶段是档案信息化的成熟期。档案信息化由重视过程信息化开始重视档案信息资源的开发利用。此时,档案部门感觉到现有电子档案管理系统在档案信息资源开发上的局限性,着手对原有信息系统进行优化和二次开发。

(5) 档案信息化高级阶段。这是档案信息化的未来发展方向,有些已初露端倪,有的学者称之为后保管时代,有的对档案信息化向知识化方向的发展充满信心,在未定之前我们暂时称之为档案信息化的高级发展阶段。

### 三、档案信息化各阶段特征模型

根据档案信息化各阶段特征,下面据此设计特征模型,用硬件、软件、流程、馆藏资源、开发利用等指标对档案信息化各阶段特征模型进行度量,具体见表 1-2。

表 1-2 档案信息化各阶段特征模型表

阶段	硬件	软件	流程	资源数字化	开发利用
单机应用	较少计算机	单机版或无专用软件	仅限于目录打印和查找	无	无
业务工作计算化	有少量计算机	有单机版或网络版	大多数环节通过计算机完成	目录级基本实现数字化	无
协同工作阶段	有较多数量计算机	网络版	基本流程,并与业务部门进行协同	着手全文数字化	开发利用实现信息化
数字档案馆阶段	全员计算机	架构先进的网络版	流程进行了优化	目录、全文全部实现数字化存储	开发利用全部信息化
高级阶段	全员计算机	与企业信息资源集成	与企业主营业务集成	进一步优化,协同	实现知识化

## 四、档案信息化发展阶段划分的意义

档案信息化阶段划分是对档案信息化工作发展过程的总结,基本反映了企业档案信息化发展的规律,其基本思想对于档案信息化建设具有指导意义,特别是在制定数字档案馆建设规划,档案部门首先应明确组织当前处于哪一个成长阶段,进而根据该阶段特征来指导数字档案馆建设工作。

这个发展阶段论也为区分当前的数字档案馆建设与十年前的数字档案馆建设提供了依据。据报道,早在2004年当时的江苏电力公司就宣称建成了数字化档案馆,随后相继有不少机构宣称建成了数字档案馆。2007年,航空工业第一集团公司宣称开始建设数字档案馆,开展数字档案馆试点建设。

近几年,在国家档案局的推动下,企业建设数字档案馆的积极性空前高涨,国家档案局也及时制定并发布了《企业数字档案馆(室)建设指南》,指导并规范此项工作。那么,当前的数字档案馆与十几年前的数字档案馆区别在哪?通过对诺兰模型的分析我们可以将数字档案馆的发展划分为辅助人工型、信息管理型和协同型三个阶段。各个阶段管理的重点分别对应目录管理、数字化全文管理和电子档案管理。从管理的重点来看,当前建设的数字档案馆属于协同型数字档案馆,如果把目录管理、数字化全文管理定义为数字档案馆1.0版和2.0版的话,那么现在建设的数字档案馆则是3.0版,它具有如下特点。

(1) 管理重点由数字化全文转为原生性电子档案管理。

(2) 电子档案管理系统不再是一个管理信息的系统，而是一个协同工作系统。

(3) 电子档案管理系统的功能有了较大提升。

(4) 大量智能技术在数字档案馆中得以应用。

(5) 知识管理技术在数字档案馆中得到重视和应用。



## 第二章 企业数字档案馆建设规划

### 第一节 数字档案馆规划原理

数字档案馆本质是信息系统，企业数字档案馆本质是企业信息系统的一种。因此，企业数字档案馆的规划本质也是企业信息系统的规划，可以借鉴信息系统规划的原理。本章先叙述信息系统规划原理，再叙述数字档案馆的规划。

#### 一、企业信息系统规划原理

企业信息系统的规划是关系企业信息系统长远发展的计划，是企业战略发展规划以及企业信息化战略发展规划的一个重要组成部分，这不仅因为企业信息系统建设是一项历时长、耗资巨大、技术复杂且又内外交叉的工程，更因为信息已成为企业的生命线，信息系统和企业的运营方式、文化习惯息息相关。

一个有效的战略规划可以使信息资源得到合理地分配和利用，可以促进信息系统的深化应用，可以使信息系统与用户建立良好的关系，可以节约建设成本。一个好的规划还可以作为一个标准，有利于对信息系统建设工作进行考核，明确建设的方向，调动有关人员的积极性。对信息系统进行规划的过程也是对业务工作进行回顾的过程，它可以使企业领导及有关人员在推进当前工作的同时“回头看”，发现问题和不足，促进整改和提升。总之，信息系统规划对企业来说非常重要，应大力提倡和推广。

企业信息系统规划包含的内容丰富，涉及面广，绝不仅是“花钱买机器”的一个简单规划。它从企业的总目标到各部门的分目标，从各部门的政策和计划到信息部门的活动与发展，均有所涉及。一个好的规划在内容上应包括：管理信息系统的目标、约束和计划指标的分析，信息系统的功能结构，信息系统的组织、人员、管理和运行，以及信息系统效益分析和实施计划等方面。做好

该规划，还应通过以下流程逐步实现，如图 2-1 所示。

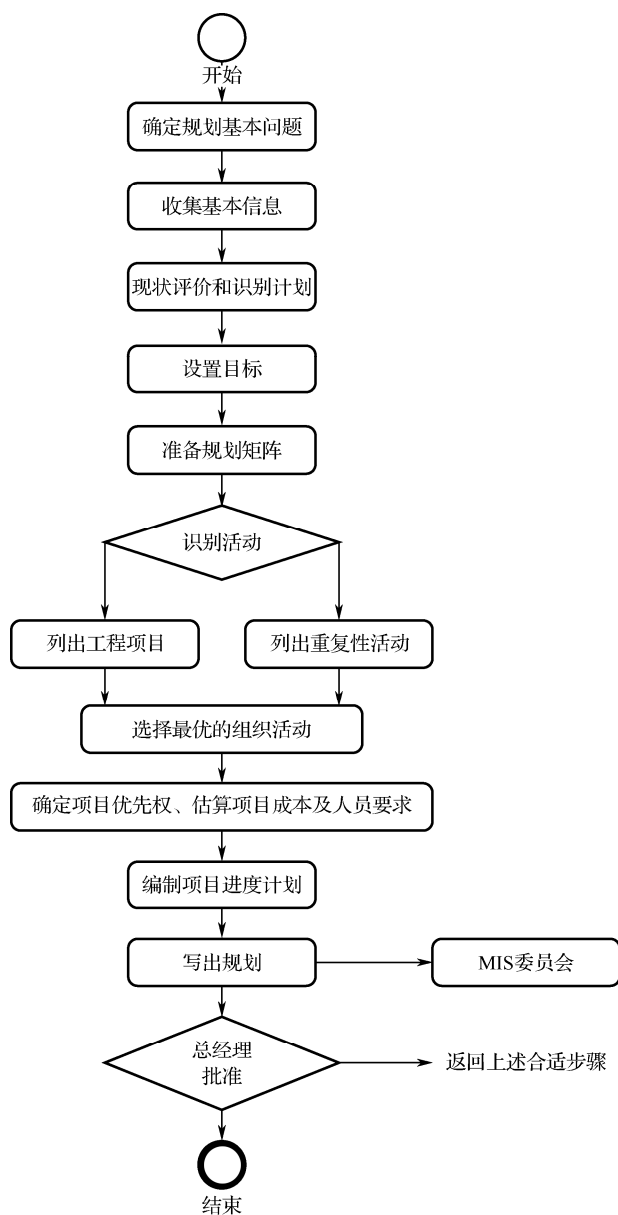


图 2-1 信息系统规划流程图

第 1 步，确定规划基本问题，包括规划的年限、规划的方法，确定集中式还是分散式的规划，是进取还是保守的规划。

第2步，收集初始信息，包括从同类企业、本企业各级管理人员、各信息系统委员会、各种文件、书籍、杂志等收集信息。

第3步，对现有状态的评价和识别计划约束，包括目标、系统开发方法对规划活动的影响、现有硬件及其质量、信息技术人员及其经验，现有信息系统运行和控制情况，资金、安全措施、标准、中期和长期优先顺序、外部和内部关系。

第4步，设置目标。此步骤应由企业主要负责人和信息化负责人执行，包括所规划的信息系统服务的质量范围、政策、机构和人员等。

第5步，准备规划矩阵。规划矩阵是指信息系统规划内容及其相互关系。这一关系确定后，信息系统规划的各项内容及其实现的先后顺序也就得以确定。

第6~9步，识别规划矩阵所列各项活动，是一次性的、工程项目性的活动，还是重复性的、经常进行的活动。由于资源有限，所有的项目不可能同时进行，只有选择一些最有效益的项目先进行，同时要正确分配工程类项目和日常重复类项目比例，正确分配风险大的项目和风险小的项目的比例。

第10步，确定项目优先权、估算项目成本以及人员要求。

第11步，编制项目实施的进度计划。

第12步，将规划书写成文，在此过程中应多次与有关人员交换意见，包括信息化工作负责人、信息技术部门工作人员、信息系统使用人员等，并不断修改完善。

第13步，将规划文本报请总经理批准生效，并宣告此规划任务完成。

## 二、档案信息化规划步骤

前面对企业信息化规划的步骤进行了研究，共分13个步骤完成。一个规划形成的过程实际上就是一个寻找发展机会，确定目标，并对目标进行论证的过程。规划的制定可分为两个比较明显的阶段，第一阶段是目标确定阶段，第二个阶段是任务分解阶段。这完全是一个利用现有论据进行结论证明的过程，特别是目标确定阶段。制定规划最核心的任务是确定发展目标，但是目标不能凭空臆测出来，它必须建立在有充分证据支撑的基础之上，即通过分析现状，分析内外部环境，分析外部机会与威胁，分析内部优势与劣势，分析标杆等的基础上形成一个可行的发展目标。前期每一个环节的工作都对形成发展目标提出

依据。

在第一章定义中论述过，数字档案馆实质就是一个信息系统，因此数字档案馆建设规划本质上也是档案信息化规划。因此，数字档案馆规划是职能信息化规划，不一定要严格按照信息化规划的步骤来执行，但主要步骤必须照此进行，作者根据自己的工作经验，结合信息化规划的理论和方法，将数字档案馆规划工作分解如图 2-2 所示的 10 个步骤。

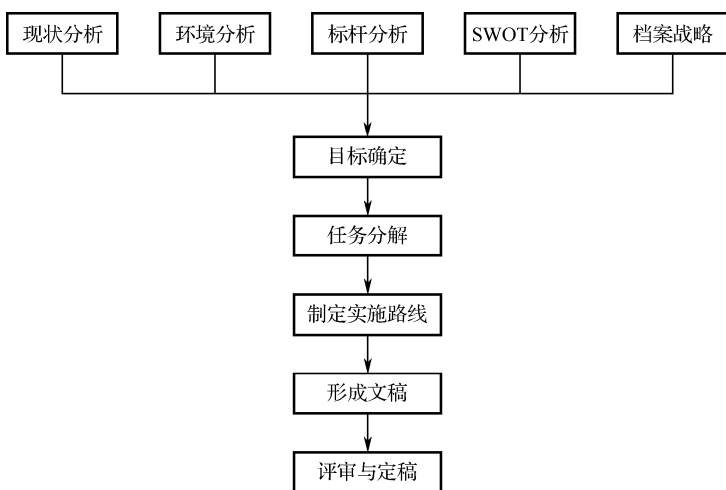


图 2-2 数字档案馆规划流程图

## 第二节 数字档案馆规划的主要内容

### 一、现状分析

根据规划的原则与步骤，规划的第一步是对工作现状进行分析，通过现状分析找出差距和弱项。数字档案馆现状分析主要从以下方面进行。

(1) 分析本企业档案信息化整体水平。即企业档案信息化在全国、本行业中的地位，是处于领先地位还是处于中间位置，或是处于靠后位置，动员状况如何，必须用定性的语句进行表述。如可将本企业档案信息化整体水平表述为“经过上一个五年的工作，我公司档案信息化获得了较好的发展，综合水平走在

中央企业前列，档案信息化较好地支持了档案工作的开展”。

(2) 分析企业档案数据库建设情况。即以档案数字化率来衡量，包括档案目录的数字化率和全文档案的数字化率，以及档案数据的数量。

(3) 分析企业档案信息化硬件配置情况。按照网络覆盖率、服务器配置、计算机台数、输入设备、加工设备、输出设备的框架进行分析。

(4) 分析企业档案管理信息系统实施情况。一般从档案信息系统功能的完备性、架构是否先进等角度进行分析。

(5) 分析档案信息化标准规范编制情况

档案信息化标准规范编制情况可从标准的覆盖程度进行分析，如果标准规范覆盖了档案信息化工作的每一个环节和每一类档案的管理，则覆盖范围较好，否则只能定位为标准规范不全。分析标准规范编制现状的另一个指标是先进性，即标准规范是否充分采用了最先进的管理理念和技术。

(6) 分析档案信息化安全保密情况

对档案信息化安全保密现状分析主要从人防、物防、技防三个方面展开。人防主要是指档案信息化安全保密是否职责到位；物防主要是指载体保管设备的配备情况；技防主要是指安全保密技术的运用情况。

(7) 分析档案信息化人员情况

从职称结构、专业结构、年龄结构、学历结构等角度分析档案信息化岗位上人员的数量，计算出比例。

(8) 分析档案信息化应用效果

采用定量与定性分析的方法对档案信息化应用取得的效果进行描述，包括员工获取档案信息的便捷程度、在线检索档案的人员数、档案利用实例等。

(9) 分析档案信息化存在的问题与不足

这是任何一个战略规划必需考虑的问题，只有找出不足才能发现短板，确定下一期战略规划的重点。

## 二、环境分析

环境决定着战略实现，因此，在制定战略前必须进行工作所处的环境分析。根据企业规划理论，规划环境分析包括宏观环境分析和产业环境分析。宏观环境分析包括政治法律因素、经济因素、技术因素和社会人文因素，行业环境包

括竞争者、潜在竞争者、供应商、客户、替代品五方面的因素。但是由于档案信息化规划是一种职能层面的信息化规划，没有必要对环境的各个要求进行分析，只需要对影响数字档案馆建设的因素进行分析，主要关注以下几个方面。

### 1. 档案工作发展大趋势

档案工作发展大趋势是指档案工作发展的宏观趋势，这个趋势决定了单个企业数字档案馆建设是否具有发展的氛围。

### 2. 企业发展大趋势

数字档案馆归根到底是企业服务的，因此，在对数字档案馆建设内外环境进行分析时，必须分析企业发展的大趋势，分析本企业是处于上升期，还是处于稳定期，还是处于衰退期，这个趋势决定了数字档案馆建设能否获取必要的资源支持，同时也决定了企业对数字档案馆是否有需求。

### 3. 企业信息化发展趋势

企业信息化发展的趋势对企业数字档案馆建设起着关键作用，必须予以关注。一般来说，信息化处于上升期的企业对数字档案馆关注较少，信息化处于稳定期的企业由于大量数据和电子文件的产生，对数字档案馆建设需求较迫切。但信息化处于上升期的企业一般会将信息化建设的重点放在主营业务上，对数字档案馆建设不会给予资源，当然这也不利于企业信息化的发展。

### 4. 企业档案工作大趋势

企业档案工作大趋势对档案信息化有着巨大的影响，档案工作根据其发展水平可划分为起步阶段、规范化阶段、信息化阶段、知识化阶段，档案工作处于起步阶段和规范化阶段的企业将很难有精力投入到档案信息化中，因此，在进行档案信息化规划时必须将本企业档案工作所处的阶段清楚定位。

## 三、标杆分析

标杆分析是选取若干数字档案馆建设或档案信息化发展较好的组织或企

业，分析其档案信息化水平较高的特征，用定性或定量的指标进行描述。标杆的作用是为下一步确定发展目标提供参照，如可将本书第十二章提供的案例作为标杆。

## 四、SWOT 分析

SWOT 分析是通过 SWOT 矩阵对企业外部环境和内部条件进行分析，从而寻找二者最佳可行战略组合的一种工具。它在数字档案馆建设中的应用是采用 SWOT 分析对企业档案部门和档案工作所面临的外部环境和内部条件进行分析，寻找二者最佳组合后的可行战略目标。在这里 S 代表档案部门的“长处”或“优势”(Strengths)；W 是档案部门的“弱点”或“劣势”(Weaknesses)；O 代表档案部门外部环境中存在的“机会”(Opportunities)；T 代表档案部门外部环境所构成的“威胁”(Threats)。进行 SWOT 分析，一般要经过下列步骤。

(1) 进行档案部门外部环境分析，列出对档案部门来说外部环境中存在的发展机遇(O)和威胁(T)。

(2) 进行档案部门内部环境分析，列出档案部门目前所具有的长处(S)和弱点(W)。

(3) 绘制 SWOT 矩阵。这是一个以外部环境中的机会和威胁为一方，以档案部门内部环境中的长处和弱点为另一方的二维矩阵。

在这个矩阵中，有四个象限或四种 SWOT 组合。它们分别是长处—机会(SO)组合、长处—威胁(ST)组合、弱点—机会(WO)组合、弱点—威胁(WT)组合。

(4) 进行组合分析。对每一种外部环境与内部条件的组合进行分析。

## 五、企业档案事业战略和企业信息化战略继承

企业数字档案馆建设规划是企业档案事业战略和企业信息化战略的重要组成部分，必须服从并服务于这两个规划，因此数字档案馆建设规划必须继承企业档案事业战略规划和企业信息化战略的思想，着眼于档案战略和信息化战略的落实与深化。

## 六、目标确定

目标确定，即确定数字档案馆建设的目标，目标要定量与定性相结合，一般用一段话进行表述。首先用一句定性的语句描述数字档案馆要达到的总体目标，然后分别从各个方面描述数字档案馆建设要达到的目标。如总体目标一般可描述为“建设具有国内先进水平的数字档案馆”、“建成功能齐全、满足公司业务发展需要的数字档案馆”。各子目标可以表述为“电子文件归档率达到 60%，传统载体档案数字化率达到 90%”等。

## 七、数字档案馆总体技术架构设计

为了清楚地表示数字档案馆设计的成果，需要用相应的架构图来表示。在信息系统规划中，用来表示信息系统规划设计成果的就是各种架构图，用得较多的有纵向架构和横向架构，其中横向架构有应用架构、程序架构和数据架构，纵向架构是指技术架构。鉴于数字档案馆本身的特点及本书阅读对象考虑，这里我们主要叙述其技术架构。

数字档案馆的技术架构是其纵向架构，是指数字档案馆系统的应用、应用平台、基础设施之间的关系。技术架构设计是档案数字档案馆规划中的重要内容，其设计是否科学决定了数字档案馆建设成果的科学性和适用性。数字档案馆技术架构以层次化结构涉及数字档案馆建设的各个方面，每一个层次由许多功能模块组成，每一功能模块又可分为更细的层次。数字档案馆总体架构还包括其下的信息技术应用架构和数据资源架构。图 2-3 是某企业数字档案馆的技术架构。

## 八、电子档案管理系统平台设计

系统平台是档案管理信息系统开发与应用的基础，它的设计包括计算机处理方式、网络结构设计、网络操作系统的选择、数据库管理系统的选择和软硬件选择与设计等。



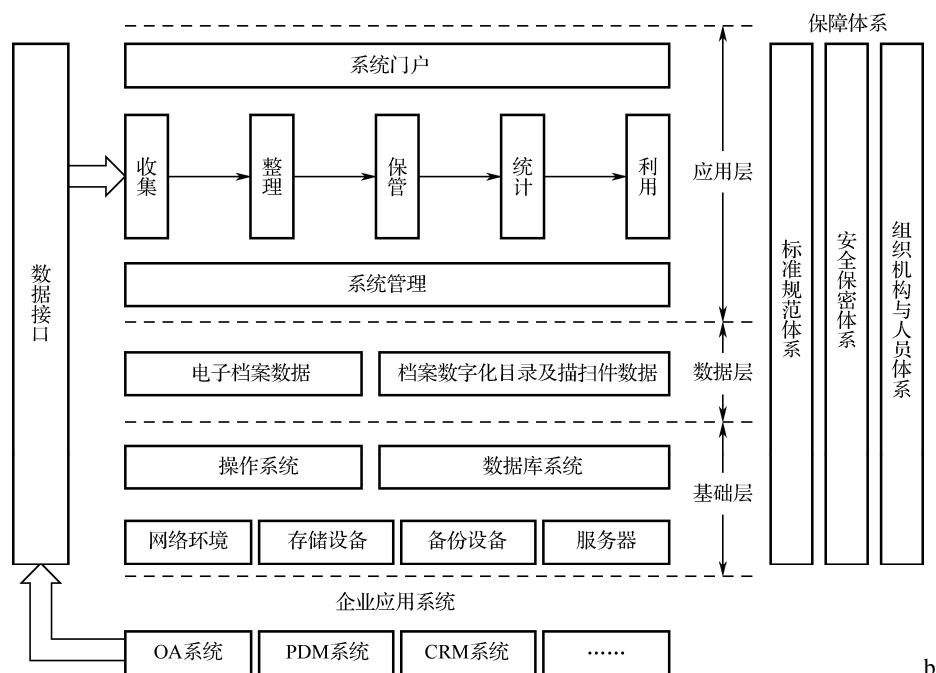


图 2-3 某企业数字档案馆的技术架构图

(1) 选择系统平台

(2) 计算机处理方式的选择和设计

计算机处理方式可以根据系统功能、业务处理特点、性能/价格比等因素，选择批处理、联机实时处理、联机成批处理和分布式处理等方式。

档案管理信息系统一般采用联机实时处理、联机成批处理和分布式处理三种方式混合使用。

(3) 计算机网络系统的设计

计算机网络系统参见第四章有关内容。

(4) 数据库管理系统的选择。

电子档案管理系统的数据库系统的选择大多基于成本、企业其他系统数据库使用现状等因素。单从成本费用考虑，电子档案管理系统数据库系统一般选用 Microsoft 的 SQL Server 数据库，从性能方面考虑可选用 Oracl、IBM DBII 等数据库，但由于系统集成需要和避免重复采购、运维成本等考虑，很多企业会选择与企业现有系统相同的数据库系统。如有的选择与 PDM 系统相同的 Oracl，或是选择与 OA 系统相同的数据库。

#### （5）软硬件选择

硬件的选择原则是：选择技术上成熟可靠的标准系列机型；处理速度快；数据存储容量大；具有良好的兼容性、可扩充性与可维修性；有良好的性能/价格比；厂家或供应商的技术服务与售后服务好；操作方便；在一定时间保持一定的先进性的硬件。

软件选择的原则包括：操作系统、汉字系统、设计语言和应用包等软件的选择。

## 九、确定重点任务

确定重点是数字档案馆规划的重点工作，即根据确定的目标、总体架构等确定为实现目标需要开展的工作和完成的任务。数字档案馆建设的重点任务一般按照数字档案馆组成架构分别叙述即可。一般来说，数字档案馆建设的有以下重点任务：

- （1）基础业务完善；
- （2）制度标准建设；
- （3）硬件购置；
- （4）软件开发；
- （5）数字档案资源建设；
- （6）安全保密体系建设。

## 十、制定规划路线

规划路线是将重点任务进行分解，设计完成重点任务，实现规划目标的途径。一般是将跨年度或重大的任务分解为若干个阶段或若干项子任务，对每一个阶段的任务或子任务分解到相应的年度，一般细化到季度。与此同时，在分解任务时还需确定实现规划目标所需的预算，计算年度投入预算和总预算。

## 十一、规划任务的外包

制定规划都必须遵循前面的流程，但如何利用这个流程则有不同的做法，这决定于规划制定方式。制定规划当前主要有两种方式，一种是组织内部自行组织人员制定，另一种是将数字档案馆建设规划外包给专业的咨询公司，两种工作方式有很大的不同。

数字档案馆建设规划外包则将规划编制工作外包给具有一定实力的咨询公司完成，企业付一定的咨询费用。中央企业或大型集团企业的数字档案馆建设规划大多采用咨询外包形式。

咨询外包方式工作流程与一般管理咨询或战略咨询相同，主要包括以下工作。

(1) 与咨询公司面谈，提出初步需求。

(2) 咨询公司提出《项目建议书》，并进行《项目建议书》的宣讲。《项目建议书》是管理咨询项目甲、乙双方进行第一轮沟通的工具，是咨询公司根据对甲方初次接触而作出的对项目的理解，是选择合适咨询公司承担规划咨询任务的主要依据。《项目建议书》一般包括项目目标、规划的工具与方法、模型、工作内容、实施过程及时间节点、项目组的组成、报价或预算。

在数字档案馆建设规划外包中，价格的核算是关键，价格高了，加大档案部门成本，价格低了有可能达不到预期的质量。一般来说，核算咨询公司价格的方法宜采用人员投入法。即根据咨询公司投入人员的情况来计算价格。咨询公司人员一般分为顾问、咨询师、助理咨询师等。顾问价格较高，咨询师低于顾问，助理咨询师最低，咨询项目总价格可用以下公式计算得出：

$$\text{顾问人数} \times \text{天数} \times \text{价格/天} + \text{咨询师} \times \text{天数} \times \text{价格/天} + \text{助理咨询师} \times \text{天数} \times \text{价格/天}$$

至于每天的价格，各咨询公司报价不一，有的上万元一天，有的几千元一天，目前国内咨询师价格一般在 3000~5000 元/天。

(3) 选定咨询公司。

大型企业咨询项目外包都必须实行比质比价，一般都必须对比多家咨询公司的项目建议书，最后选定一家对项目理解较透彻、价格相对较低的咨询公司。选定咨询公司后，需签订项目合同，通过合同约定服务的内容、价格、完成时间等。

(4) 召开项目启动会。

合同签订后，即可启动项目。项目启动前一般会召开项目启动会，标志着项目的正式启动。项目启动会参会人员可邀请双方领导、档案工作主管领导、信息化工作主管领导、档案工作部门负责人、信息化工作部门负责人、咨询公司项目组人员、档案部门项目负责人等人员参加。会议内容主要是双方领导讲话，咨询公司人员介绍项目进度安排。为了给档案工作营造较好的氛围，会议最好悬挂会标，会后撰写消息在内部媒体上发布。

(5) 咨询工作配合。

项目启动会召开完毕即标志着咨询项目正式启动，咨询公司即进入正式工作状态。此前咨询公司会要求项目方提供工作场所，项目方准备一间相对独立的会议室较为适宜。咨询公司在项目工作中要做大量的信息收集与人员访谈工作，并且需要访谈不同层次的人员，期间对工作内容还需要进行交流。这要求档案部门根据咨询公司提供的人员访谈时间安排，提前预约相关人员，使项目咨询工作顺利进行，直至项目完成。

(6) 项目验收。

项目进行到最后，会形成咨询成果，即规划思路，有的公司以文本文档的形式提供，有的咨询公司以 PPT 文档的形式提供，但不管以何种形式提供，都需要对成果进行一次验收评审。评审最好以会议的形成进行，形成评审意见和结论。

### 第三节 规划文本的形成、评审和项目立项

#### 一、规划文本的形成

规划的最终结果是规划文本，它的结构相对比较固定，一般结构如下。

##### 1. 引语

用一段简短的话说明规划的目的和主要内容。

## 2. 档案信息化现状

有时也用“档案信息化所取得成绩和存在的问题”作标题，主要内容不外乎说明档案信息化的成绩和存在的不足，篇幅不需要太长。

## 3. SWOT 分析

按外部机遇、外部威胁、内部优势、内部劣势的顺序叙述行文，一般分为四个要点。

## 4. 档案信息化发展趋势及标杆

档案信息化发展趋势只需用简短的一段话概括即可。标杆分析可用二维表格方式呈现，将标杆按照档案信息化发展现状的项目一一列出即可。

## 5. 发展目标

用一段话将档案信息化工作要达到的目标描述出来。首先用一句定性描述的语句提出总体发展水平，然后分别从不同的方面给予具体可量化的目标描述。

## 6. 重点任务

重点是规划文本的主体内容，可按条、段方式逐一叙述。为增加文采，可将内容概括成数字组合，如“124”任务，即一个系统开发，两个数据库建设，夯实三个方面的基础，共七项任务，内容易记，而且有文采。

## 7. 实施计划

实施计划一般分为起步（或准备）、重点实施、全面完成三个阶段，三个阶段的时间分配一般是重点实施阶段较长，而起步和全面完成阶段稍短，以 3: 4: 3 的比例分配较为合理，不过也必拘泥于此，需根据实际情况进行时间分配。接下来就要逐一对重点任务进行分解，确定大致的时间进度，将进度逐一分配到各个阶段中，列出年度工作计划。

## 8. 主要措施

按点写出实现规划目标、完成规划任务需要采取的措施。措施不能太少，但也不能太多，一般以五至七条措施为宜。

## 二、规划文本的评审

数字档案馆建设涉及技术路径的选择，为了保证项目的顺利进行，数字档案馆建设方案制订完成后，应聘请内外部专家，对建设方案进行可行性论证。专家一般以 5 至 7 人为宜，应聘请精通数字档案馆建设方面的档案专家和信息化专家。评审中，应吸收专家提出的意见和建议，对设计方案进行必要的修改。

方案经过评审后，数字档案馆的进度、技术路线、经费等已基本确定，接下来就是落实经费。项目在本企业立项，只有通过立项才能将所需的经费纳入预算。一般来说，数字档案馆建设项目所需经费应纳入信息化建设项目预算，或是固定资产投资预算，确保所需经费得到落实。

## 第三章 企业数字档案馆建设实施与运维

前一章叙述了企业数字档案馆建设规划，最终成果是规划报告。有了规划报告，企业数字档案馆建设有了依据，一旦立项并解决了经费，就可进入启动实施阶段。本章就叙述企业数字档案馆的实施工作。企业数字档案馆建设项目实施，包括标准制度制定、硬件配置、电子档案管理系统实施、传统载体档案数字化和电子文件归档实施、安全保密体系建设、项目验收与评价等。其中，安全保密体系建设应贯穿整个项目实施的全过程。

### 第一节 实施准备

#### 一、制定实施方案

明确实施目标，详尽设计技术方案，明确团队分工及协调机制，详细分解任务并细化工作进度。其中详尽设计技术方案是重点，各部分详细的技术方案参见本书各有关章节。

#### 二、人才队伍建设

应通过提升现有人员业务技能和专业知识的途径，充分发挥现有人员的作用；现有人员不足的，应通过引进、组建项目组等方式补充。

#### 三、成立项目工作组织

数字档案馆建设是一项系统性工程，需要档案部门、信息技术部门、业务

部门、软硬件厂商等多方人员参与。为了便于组织，需要以项目组的形式将各方人员组织协调起来，因此，数字档案馆建设初期的首要任务就是成立项目组。

项目组织可根据数字档案馆建设规模的不同而采用不同的组织方式。如采用集中部署或统一建设集团型数字档案馆的，应分别成立数字档案馆建设项目领导小组、工作组和推进小组。

### 1. 领导小组

领导小组设组长一名，副组长若干名。其中，组长可由总经理或总工程师等主要或分管领导担任，副组长可由档案部门、信息技术部门等部门负责人担任。小组成员应包括：档案部门、信息技术部门、文书处理部门（办公室、综合部等）以及经营计划、科技管理、基层设计、工艺、生产等有关部门负责人。

主要职责有：对数字档案馆建设实施方案进行决策；对工作组制定的电子档案管理系统需求目标进行决策；对工作组制定的电子档案管理系统工作网络计划进行决策；检查督促各单位各部门的实施情况等。

### 2. 工作组

工作组设组长一人，副组长若干名。其中，组长可由档案部门负责人或信息技术部门负责人担任或兼任，副组长可由上述两部门的业务技术骨干担任。成员应包括档案部门、信息技术部门、文书处理部门等的业务或技术人员。

主要职责有：负责制订规划方案；制定电子档案管理系统实施、运行、管理的规章制度和标准规范；组织对现有业务流程进行分析并提出电子档案管理系统需求分析；组织制定并实施档案资源综合管理工作的网络计划；负责组织电子档案管理系统的操作培训；负责电子档案管理系统及其依赖环境的安装、调试及相关培训工作等。

### 3. 各职能部门和下属单位推进小组

由于数字档案馆建设涉及每一个部门和每一个下属企业的所有员工，为使项目的实施落实到每一个部门和每一个下属企业，各部门各下属企业应组织相应的推进小组。推进小组由部门（或下属单位）一把手或负责档案工作的负责



人任组长，部门专（兼）职档案员任副组长，有关人员任成员。如果是下属单位，档案部门负责人任副组长，其他各室（班、组）人员任副组长或成员。其职能是：提出本部门电子档案管理系统的需求；制订电子档案管理系统在本部门实施的方案及组织实施。

## 四、建立项目组运作机制

各项目组的成立为项目实施运转奠定了组织基础，但因项目组成员来自多个部门、多个单位，具有临时性的特点，项目组的运转往往成为项目实施过程中的一个难题。因此，制定一套有效的运转机制，增强项目组的凝聚力，是十分必要的。一般来说，这套机制可以从以下几个方面考虑。

（1）制定科学的项目计划作为协调各小组及各成员活动的核心。

（2）明确各小组及小组成员的分工与职责以防止推诿扯皮。

（3）制定严格的管理制度，对规范小组活动十分必要。

（4）定期召开小组会议。一般来说，领导小组会至少每半月应召开一次，推进小组会至少每周应召开一次。在实施的关键时段，如系统试运行阶段，每天应召开推进小组会议。

（5）构建小组信息交流渠道和平台。如可建立一个临时性的网站，发布或交流实施进度信息，或者定期编发实施简报，定期向领导小组和各部门推进小组通报信息。

## 第二节 项目实施

### 一、项目实施启动

项目实施启动即以召开启动会的形式正式启动项目的实施工作。项目启动会应由数字档案馆建设领导小组、工作组及推进小组成员参加，必要时还应邀请上级档案部门和信息技术部门负责人及有关人员参加，会议议程一般包括以下方面。

(1) 领导小组组长讲话，强调数字档案馆建设的重要性，对数字档案馆的建设提出要求，对实施工作提出要求等。

(2) 由领导小组副组长或工作组组长介绍实施工作方案。

(3) 上级档案部门领导讲话。

(4) 宣布启动。

## 二、基础业务完善与规范化

数字档案馆建设是以现有的档案工作为基础开展的，在项目启动阶段，需对本企业档案管理基础业务进行完善与规范。主要包括：健全或完善档案管理制度、理顺档案管理体制机制、实现档案工作的集中统一、规范档案管理业务流程及各类卷夹、表格、单据等。

## 三、制度规范制定

根据项目进展情况有步骤地制定制度规范。在数字档案馆建设中应优先制定有关工作标准规范，应优先采用国内成熟标准规范作为本企业数字档案馆建设标准规范，国家标准规范不能满足需求或无相应标准规范的应自主编制，所形成的标准规范应定期更新。在数字档案馆运行前，应完成所有制度规范的制定，以使运行规范化。

## 四、基础设施配备

鼓励企业依托现有信息化基础设施配备数字档案馆建设和运行所需硬件，以节约投资成本，提高人力、物力集成的管理效率。对于现有基础设施无法满足系统使用人数、安全保密需要以及未来3至5年数据增长时，应予以配备。同时，硬件配备还要根据电子档案管理系统的部署方式进行配备，如全集中式、半集中式、分散式等。

## 五、电子档案管理系统实施

电子档案管理系统实施工作包括系统开发、基础数据准备、系统测试、人员培训、系统切换和试运行、系统评价与验收等工作，具体参见第七章和第八章。

## 六、传统载体档案数字化实施

按照“重要性、常用性、急用性、抢救性、珍贵性”五原则，并结合实际开展传统载体档案数字化工作，将重要的、常用的、急用的、自然损坏老化严重的和珍贵的档案优先数字化。传统载体档案数字化应以本企业自主完成为主，自主完成有困难的可采用外包，或外包与自主相结合方式进行。档案数字化采用外包的，应按照《国家档案局办公室关于印发<档案数字化外包安全管理规范>的通知》（档办发〔2017〕7号）要求做好安全管理工作。传统载体档案数字化内容有如下几方面。

（1）经营管理、生产技术管理、行政管理、党群工作等管理类档案应按照先近后远的原则，将最近形成的档案先行数字化，以满足频繁利用的需要。对于即将到期的短期档案或10年期档案可不进行数字化。

（2）产品生产和服务业务类档案可按照在研和在产核心产品档案或主营业务的业务档案、在研和在产非核心产品档案或非核心业务档案、停产产品档案的顺序进行数字化。

（3）科研档案应按照获奖等级、项目下达部门的层级顺序进行数字化，即先数字化获得国家级奖励的，再数字化获得省级、部委级奖励的，最后数字化获得企业级奖励的；对于未获奖的科研档案按照国家科技项目、省部级科研项目、企业级科研项目的顺序进行数字化。

（4）基本建设档案可按照建筑物的功能、用途的重要程度来确定数字化顺序。可将生产、科研建筑物档案先数字化，其次是办公建筑，最后是生活建筑。基本建设档案数字化的内容也要进行鉴选，一般来说可行性研究阶段、设计阶段、项目管理、施工阶段的竣工图和质量管埋、竣工验收等阶段形成的档案利用较多，应进行数字化，其他阶段形成的档案可选择性地进行数字化。

(5) 设备仪器档案可按设备的价值、用途等标准来确定其档案数字化的顺序。

(6) 会计档案由于数量较大,一般越晚形成档案的查阅利用越频繁,可按照形成时间由近及远进行数字化,即将最近形成的档案先数字化,年代久远的后数字化。

(7) 人事档案的数字化,档案部门应与人力资源管理部门充分协调确定后按照国家组织、人事部门的有关规定开展数字化。

(8) 在企业数字档案馆建设过程中,应制定专题数据库建设规划,以需求为导向,选择利用率高、数量稀少、内容重要的档案资源,以文字、图片、音视频等多种方式展示利用。

各类档案数字化方法参见第五章。

## 七、电子文件归档实施

电子文件归档是数字档案馆建设中的重点和难点,在数字档案馆建设期间一般不全面开花,而且选择一到两个系统开展电子文件归档工作,一般按以下步骤开展。

(1) 确定数字档案馆建设期间纳入归档的业务系统。

一般应选择运行时间长、形成电子文件较多的主营业务系统作为建设期间的归档对象。如制造企业可选择 ERP、PDM、财务管理等系统形成的电子文件作为归档对象;服务企业可选择财务管理系统、电子邮件系统、电子商务系统形成的电子文件作为归档对象。这些工作需在数字档案馆建设规划时完成,纳入规划。

(2) 分析待归档电子文件形成系统电子文件的内容、特点(存储格式)、数量等。

(3) 制定待归档电子文件形成系统电子文件归档范围。

要详细列出需归档电子文件的内容、原格式、存档格式、保管期限等。

(4) 设计元数据项及捕获节点。

(5) 设计归档接口方案。

将接口落实到电子档案管理系统需求中,包括元数据需求。

(6) 归档接口开发。

与电子档案管理系统开发、实施一起，将归档接口方案在电子档案管理系统中实现。有的还需要形成系统并进行必要的开发，有关接口开发工作参见第六、七、八章有关内容。

(7) 归档测试。

与电子档案管理系统开发、实施一起，对归档接口进行测试，测试时要有了一定的数据量。

(8) 正式归档。

对接口测试可行后，进行正式归档。

有关电子文件归档的详细内容见第六章。

## 第三节 项目实施验收

### 一、验收内容和目的

数字档案馆建设项目完成时，应对实施工作进行总结和评估，即验收。主要是对数字档案馆建设项目任务完成情况、是否达到当初设计的效果、技术水平及其先进性等方面进行评价和总结，找出差距，切实保证数字档案馆建设的效果。

### 二、验收依据

#### 1. 数字档案馆建设规划

建设规划是数字档案馆建设的重要依据，建设完成的数字档案馆是否达到规划确定的任务目标是评价建设任务是否完成的重要依据，因此，要把数字档案馆建设规划作为验收数字档案馆建设工作的重要依据。

#### 2. 数字档案馆实施方案

依据建设规划制定的实施方案，是具体指导数字档案馆建设步骤的重要文

件，它规定了数字档案馆建设的进度、经费预算、具体技术指标等，因此，也是评价数字档案馆建设工作的重要依据。

### 3. 国家有关规章制度和标准规范

国家有关数字档案馆建设的规章制度和标准规范一般在制定规划和实施方案时融合到有关要求中；但由于规章制度和标准规范更新较频繁，有些在规划和实施时采用的规章制度和标准规范在建设完成时却已出台新的版本或新的内容，因此，应在验收时作为依据纳入其中，使数字档案馆建设符合国家有关要求，尤其是安全保密方面的要求。

## 三、验收方式

应聘请内部、外部专家组成专家组，采用会议的形式进行验收。专家组一般有 5 到 7 人具有中级以上技术职务的人员担任，其中具有高级以上技术职务的专家人数应在 60% 以上，所聘专家的专业应具有广泛的代表性，档案专业、信息技术专业、企业主营业务专业等三个方面的专家都应有。验收会议由企业负责人主持，项目实施负责人介绍有关情况并演示系统，专家质询，最后形成验收意见。

## 四、验收意见

数字档案馆建设项目验收意见应包括以下内容：验收会议召开的时间、地点、主要内容，数字档案馆建设项目任务完成情况，数字档案馆是否达到当初设计的效果，数字档案馆的技术水平及创新点，存在的问题和建议。

## 第四节 项目运维

当数字档案馆建成后，电子档案管理系统安装、转换和投入正式运行，数

字档案馆即进入了运行期。在运行期间，数字档案馆中的电子档案管理系统需要进行一系列的维护工作，包括程序维护、数据文件维护、代码维护、机器设备维护等。

随着系统复杂程度的提高，系统的维护成本逐年增加。据统计，在系统的整个生命周期中，三分之二以上的经费会用在维护上，90%的软件人员在从事系统维护工作，这些数据说明维护工作在系统整个生命周期中的重要性。现实中，有的企业花巨资购买了电子档案管理系统，安装后系统长期不能顺畅运行，这与维护工作有直接关系。

项目运维即信息系统的运行管理与维护，主要包括系统的运行和系统的维护两个部分。

## 一、信息系统运行管理的原理

### 1. 信息系统运行管理的目标

信息系统运行管理的目标，就是对信息系统的运行进行实时控制，记录其运行状态，进行必要的修改与扩充，以使信息系统运行能真正符合管理决策的需要。

### 2. 信息系统运行管理的内容

#### （1）信息系统日常运行管理

##### ① 例行的信息处理及服务工作监控。

常见的工作包括：例行的数据更新、统计分析、报表生成、数据复制及保，以及与外界的定期数据交流等。

##### ② 计算机自身的运行与维护。

这里所说的计算机系统自身的运行与维护，主要是指硬件的运行管理，主要包括：设备的使用管理，定期检修，备品配件的准备和使用，各种消耗性材料的使用和管理，以及电源和工作环境的管理等。

##### ③ 系统的安全管理。

系统的安全管理，是指为防止系统外部对系统资源不合法的使用和访问，保证系统的硬件、软件和数据不因偶然因素或人为因素而遭到破坏、泄露、修

改或复制。系统的安全性即体现在保密性、可控制性、可审查性和抗击性四个方面。

## （2）信息系统运行情况的记录

系统的运行情况对系统管理评价是十分重要且十分宝贵的资料，这些资料对总结系统运行经验、提高信息处理水平具有十分重要的意义。在系统运行过程中，需要收集和积累的资料包括以下方面。

### ① 工作数据记录。

每天（周、月）提供报表的数量、每天（周、月）输入数据的数量、系统中积累的数据量、修改的程序数量、数据使用的频率以及满足用户临时需求的数量等反映系统的工作负担、所提供的信息服务的规模以及计算机应用系统功能的最基本数据。

### ② 工作效率数据。

即系统为了完成所规定的工作，占用了多少人力、物力及时间。

### ③ 系统所提供信息服务的质量。

包括使用者对提供信息的方式是否满意，对所提供信息的精确程度是否符合要求，对信息提供的及时性和临时性提出的信息需求是否得到满足等。

### ④ 系统的维护修改情况记录。

包括维护工作的内容、情况、时间和执行人员等。

### ⑤ 系统的故障情况。

需要记录的信息有：故障发生的时间、故障的现象、故障发生时的工作环境、处理的方法、处理的结果、处理人员、善后措施及原因分析。需要注意的是，这里所说的故障不仅指计算机本身的故障，而是对整个信息系统来说的，它还包括软件、操作、配置等故障。例如，由于数据收集不及时，使年度报表的生成未能按期完成，这是整个信息系统的故障，但不是计算机硬件系统的故障；再如，由于原始数据收集质量不高，使系统某项功能无法完成等。

在上述信息系统运行情况的记录中，运行数据较容易被忽视，而发生故障的情况往往记录较全。要全面掌握系统的情况，必须重视正常运行时的情况记录。

现在的信息系统一般都有日志管理功能，可自动记载自身运行情况，但是也需要有手工记录作为补充手段，因为有些情况是无法通过信息系统自身来记录的。比如，使用者的满意度、所生成报表的使用频率等，还有诸如故障发生时系统往往无法详细记录故障情况。因此，对于一套信息系统来说，必须有严



格的运行情况记录制度，并要求有关人员严格遵守和执行。

### （3）对系统运行情况进行检查与评价

信息系统在运行过程中，除了要进行大量的管理和维护工作以外，还要定期对系统的运行状况进行审计和评价。评价由系统分析员或专门的审计人员会同各类开发人员和业务部门经理共同参与，审计结果为系统的改进和扩展提供依据，系统评价一般包括以下三个方面。

- ① 系统是否达到预定目标，目标是否需要修改。
- ② 系统适应性、安全性评价。
- ③ 系统的社会效益、经济效益评价。

对系统定期进行审计与评价，实际上是衡量系统所处的状态是否满足需求。如果审计结果是系统基本适用但需要一些改进，则要做好系统的维护工作；如果审计结果确认系统已不能满足各项管理和决策需求，不能适应未来的发展，则说明该信息系统已经完结了它的生命周期，必须提出新的开发需求，开始另外一个新系统的生命周期，整个开发过程又回到系统开发的最初阶段。

## 3. 信息系统运行管理的组织与人员配置

### （1）信息系统运行管理的组织

有效地组织好信息系统运行，对于提高信息系统的运行效率具有重要意义。系统运行组织的建立是与信息系统在企业中的地位分不开的，目前，国内企业组织中负责信息系统运行管理的大多是信息中心、计算中心或信息处等信息管理部门。

### （2）人员配置

人员配置是信息系统运行成败的关键。系统运行所需人员主要包括：系统主管人员、系统操作人员、系统管理人员等。

## 二、信息系统维护的一般原理

### 1. 系统维护的内容

根据运行维护原理，一个信息系统维护包括以下主要内容。

#### （1）程序的维护

程序维护存在以下几种情况：一是根据运行记录发现程序错误，相应程序

需要更改。二是随着用户对系统熟悉程度的增加，用户对系统功能提出更高的要求，部分程序需要改进。三是系统运行环境发生变化，部分程序需要修改。

#### （2）数据文件的维护

当业务发生变化，需要建立新的数据文件，或对现有数据文件结构进行修改。

#### （3）代码的维护

随着系统运行环境的变化，旧代码不能适应新的要求，必须进行改造，制定新的代码或修改旧的代码体系。代码维护的困难主要是新代码的贯彻。

#### （4）机器、设备的维护

包括机器、设备的日常维护与管理。一旦发生故障，要立即有专人修理，保证系统的正常运行。

### 2. 系统维护的类型

因信息系统需要维护的原因不同，其维护的类型也不同，大致可分为以下四类。

#### （1）更正性维护

由于程序正确性验证尚未得到圆满解决，系统测试又不可能找出程序中的所有错误，所以，即使交付使用的系统也可能隐藏一些错误，并在一定的使用环境中暴露出来。更正性维护就是在系统使用过程中发现隐藏的错误后，为了诊断和改正这些错误而进行系统更新的活动，工作内容包括诊断问题与改正错误。

#### （2）适应性维护

为了适应外界环境的变化而增加或修改系统部分功能的维护工作。例如，新的硬件系统问世，操作系统版本更新，应用范围扩大。为了适应这些变化，信息系统需要进行维护。

#### （3）完善性维护

用户在使用系统的过程中，随着业务的发展，常常希望扩充或改进原有的功能或性能，提出新的需求。完善性维护就是指为改善系统功能或应用户的需要而增加新功能的维护工作，这类维护工作占总维护工作的绝大部分。

#### （4）预防性维护

这是主动性预防措施。对一些使用寿命较长、目前能够满足业务需要，但可能会发生变化的部分进行维护，以适应将来的修改或调整。例如，将专用报

表功能改为通用报表功能，以适应将来报表格式的变化。

### 三、数字档案馆运维工作

随着档案工作信息化应用范围的拓展，电子档案管理系统越来越庞大，投资额也越来越高，系统功能和使用也更加复杂，系统运维工作的好坏直接关系到系统的运行质量。电子档案管理系统的运行维护工作主要包括以下方面。

#### 1. 将维护事项纳入采购事项

有些电子档案管理系统的维护工作是程序性的，这类维护事项一般由软件服务商负责，有时维护工作量相当大。为了使程序性维护工作能顺利进行，防止拖延，必须将程序性维护有关条款纳入软件采购合同，明确软件服务商的维护任务、响应时间、免费周期或收费方式等。

#### 2. 要求软件服务商提供合格的开发文档

电子档案管理系统运行维护过程中需大量查阅开发文档，因此，企业在采购电子档案管理系统时，必须要求软件服务商提供完整、详备的开发文档。根据现有的商务规则，软件服务商一般只提供系统操作说明书。实践证明，仅有系统操作说明书是不能满足系统运行过程中的维护需求的。如果能提供系统的数据字典等文档，就更加方便数据文件的维护。数据字典包括数据表结构和数据表之间的关系说明，常以数据模型的方式呈现。

#### 3. 配备合适的维护人员

应根据需要配备一定数量、专业技能较强的系统维护人员，包括硬件维护人员、软件维护人员和系统管理员，他们的主要职责有以下方面。

##### （1）硬件维护人员

主要负责存储设备、网络设备、服务器、打印机、PC 等硬件维护。

##### （2）软件维护人员

主要负责代码维护的实施、数据的定期备份、系统使用指导等。

### （3）系统管理员

主要负责系统用户的增加、删除、修改、口令重置等数据维护。

当然，上述人员的配备与职责分工也不是绝对的，可以集多种角色于一人，也可以由企业信息技术部门统一安排，如硬件维护，可多业务系统硬件一并维护。

## 4. 对运维人员做好培训

运维人员充分了解系统的设计架构和各项性能、全面掌握系统的每一个细节是运维人员开展运维工作的首要前提，因此，对运维人员进行系统培训是十分必要的。培训方式可以有两种：一种是软件服务商对系统架构、性能等方面的逐一介绍，另一种是运维人员全程参与系统实施。其中，第二种方式是运维人员了解系统的最好方法，但实际工作中往往被忽视。

## 5. 制定系统运维管理制度规范

电子档案管理系统的运维必须有一套完善的制度规范，以保证其运维工作有章可循，这些制度规范应包括以下方面。

### （1）系统运维管理流程

合理的运维管理流程如图 3-1 所示。



图 3-1 运维管理流程图

### （2）系统修改事项的审批权限

根据修改对系统影响的大小确定系统维护修改事项的审批权限。

### （3）系统运维记录管理

### （4）系统运维管理办法

系统运维管理办法一般包括以下内容。

#### ① 系统运维管理的职责分工。

划分好档案部门、信息技术部门在档案管理信息系统运行中的工作职责，

特别是数据备份与恢复和系统安全保密职责，一定要明确。

#### ② 用户权限分配制度。

定义好系统运行中用户的角色，根据岗位设置，一般将系统用户定义为以下角色。

系统管理员：系统用户管理、访问控制、数据备份、状态监视。

档案管理员：档案数据输入、修改、审核、检索、统计。

兼职档案员：所在部门档案数据输入、修改、移交、检索。

普通用户：档案数据检索、利用。

#### ③ 系统操作管理。

规范系统登录、数据处理、提交、审核等工作。

#### ④ 系统异常处理与报告。

系统出现错误、宕机等情况的处理或报告流程。

#### ⑤ 信息安全与保密制度。

用户名与口令设置、防病毒措施实施等要求。

#### ⑥ 数据备份制度。

备份载体及其保管，备份周期，备份机制，数据恢复机制。

#### ⑦ 系统变更处理。

#### ⑧ 纪律与责任。

#### ⑨ 附则。

## 第四章 基础设施建设

基础设施是数字档案馆运行的支撑，它包括计算机、服务器、网络、存储及备份设备、数据库及操作系统等，企业数字档案馆的基础设施大多由企业信息技术部门统一建设，并不是数字档案馆建设的重点。本章主要对数字档案馆基础设施建设的需求进行叙述，不对其原理作详细的研究。

### 第一节 数字档案馆网络及服务器建设

#### 一、数字档案馆网络平台架构设计

数字档案馆建设中的网络平台建设，一般依托企业现有的网络平台。但是企业现有的网络平台有可能是基于日常办公需要，要满足数字档案馆运行，还需进一步进行搭建。

##### 1. 网络平台建设需求

(1) 满足电子档案管理系统服务器运行维护的需要。

要将电子档案管理系统软件、数据库、全文检索服务组件等所在的服务接入网络。

(2) 满足电子档案管理的需要。

将承担电子文件收集、整理、存储和电子档案保管等需要的终端计算机接入网络。

(3) 满足档案数字化处理的需要。

档案数字化需要专门的空间，因此，在数字档案馆建设期间需要搭建专门进行数字化处理的网络，网络接口数要能满足数字化处理的需要。

(4) 满足电子档案检索利用的需要。

满足企业正常业务对电子档案检索利用的需要，一是要将电子档案管理系统服务接入企业网络，使有档案检索利用需求的用户能方便地使用电子档案管理系统的检索功能，且网络响应速度要满足用户需要。二是必要时要搭建电子档案检索利用中心网络。电子档案利用检索中心是为了满足特定用户检索档案需要而单独开辟的区域，因此，也必须有网络覆盖，网络接口数应能满足特定用户检索档案的需要。

（5）满足智能库房的需要。

当前开展的数字档案馆建设大多有智能库房管理功能。智能库房中的档案存储环境智能化管理功能需要采集环境温湿度，对温湿度监控设备进行自动控制，在网络建设时应考虑到这方面的需求。除此之外，条码管理、出入库管理等应用都可能需要将网络布置到库房。

（6）满足安全保密的需要。

一是要建设专门用于安全保密管理的网络接口；二是网络传输线路、路由器等网络设备应符合相应的安全保密要求；三是网络拓扑结构设计应服从安全保密的要求，如相应的线路冗余度等。

## 2. 建设内容

（1）服务器接入。将安装电子档案管理系统软件的服务器、数据库服务器和文件服务器等接入网络。

（2）电子文件归档终端的接入。将兼职档案员用于电子文件归档的计算机终端接入网络。有的电子文件归档终端可能不在档案部门。

（3）数字化处理中心网络搭建。搭建数字化处置中心网络，有的数字化处理中心是一个临时场所，使用临时线路即可。

（4）电子档案利用中心。电子档案利用中心终端设备相对比较集中，但对带宽有较高的要求。因为电子档案用户要求有较好地检索体验，要求数据能在短时期内获得，且有时数据量较大，因此对带宽的要求要大于一般电子档案管理的需求。

（5）库房网络搭建。

实施智能库房管理功能的数字档案馆要在库房建设连接以下设备的网络。

- ① 建设连接采集库房温湿度数据终端的网络；
- ② 建设连接开闭空调、去湿机等设备控制终端的网络接口；

③ 建设用于连接条码枪、RFID 等管理设备的网络接口；

④ 建设用于连接出入库检测设备的网络接口。

(6) 运维管理终端接入。

包括日常运维所需的备份终端网络接口等。

### 3. 网络设计

依托企业现有网络基础架构及数字档案馆建设运行的需要，数字档案馆网络设计，如图 4-1 所示。

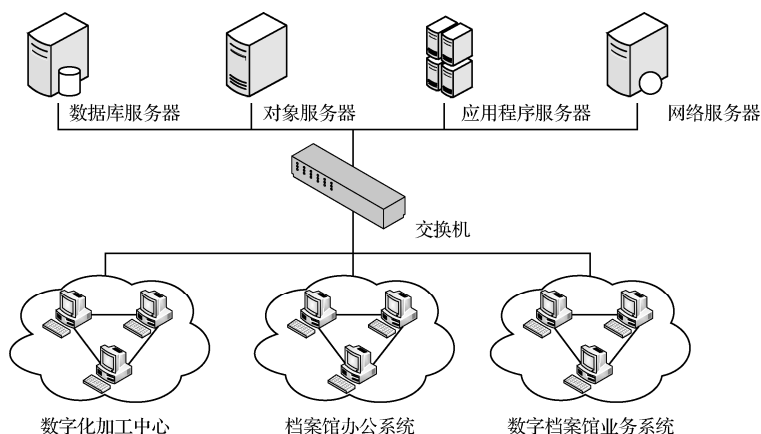


图 4-1 数字档案馆网络架构图

数字档案馆网络建设更详细的设计，如图 4-2 所示。

## 二、服务器配置选择

在数字档案馆建设过程中，经常要档案部门提出服务器需求，此时，档案部门人员经常一头雾水，所提的服务器性能太高，可能浪费经费，反之可能不能满足应用需求。服务器性能应根据并发用户数和数据量来确定，具体可参见表 4-1。



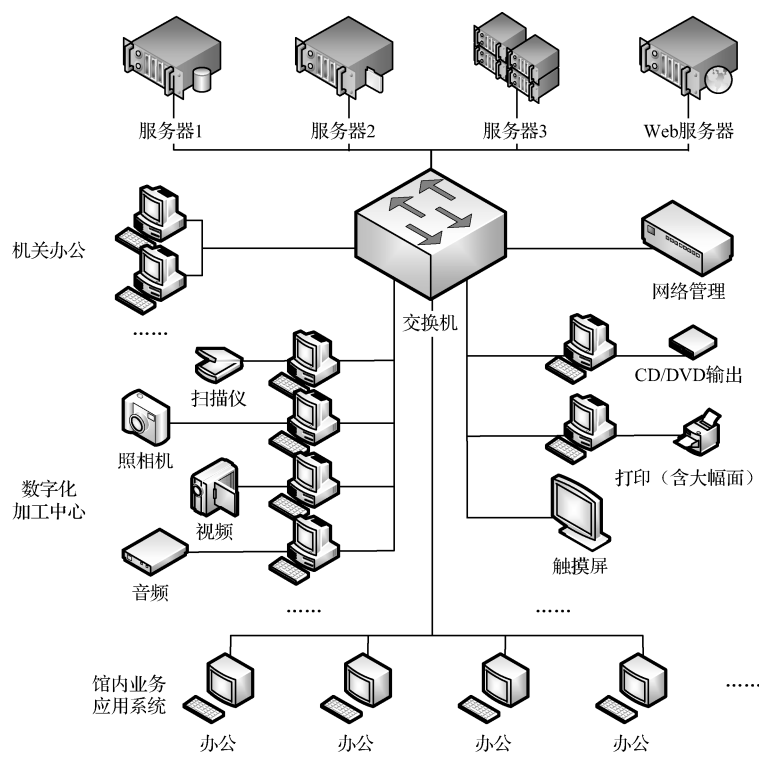


图 4-2 数字档案馆网络详图

表 4-1 服务器配备表

并发 用户数	建议服务器	CPU 的类型和数量	内存大小	硬盘大小	RAID 方式
50	一般 PC 服务器	双 Intel Xeon 2.5GHz	1~2GB	36G X4	10
100	一般部门级 PC 服务器	双 Intel Xeon MP 2.7GHz, 533MHz 前端总线	4GB	36G X6	10
200	部门级 PC 服务器	4 路 GHz 的 Intel Xeon MP 3.0GHz 处理器	4~8GB	36G X6 (磁 盘阵列)	10
300	小型机	8 路以上 SMP (1 个 4 路 MCM); 1.2GHz POWER4+	8~16GB	36G X6	10

随着云计算技术的应用，各企业一般不采用独立服务器的形式，而是采用虚拟机形式来分配其资源。虚拟服务器也称为 VPS（Virtual Private Service，虚拟专用业务）主机，简称虚拟主机。相对于真实主机而言，虚拟服务器是指采用特殊的软硬件技术把一台完整的服务器主机分成若干个主机。运行在虚拟服

务器上的操作系统和应用程序不直接控制内存、硬盘和网络端口等资源，而是由位于操作系统和应用程序之间的 VM（虚拟机）截取与硬件交互的请求，并在认为合适时机对它们进行处理。实际上是将真实的硬盘空间分成若干份，然后租给不同用户，每一台被分割的主机都具有独立的域名和 IP 地址，但共享真实主机的 CPU、RAM、操作系统、应用软件等。虚拟服务器技术是互联网服务器采用的节省服务器硬件成本的一种技术，其主要应用于 HTTP、FTP、EMAIL 等多项服务，将一台服务器的某项或者全部服务内容逻辑划分为多个服务单位，对外表现为多个服务器，从而可以充分利用服务器硬件资源。

有时单个服务器无法满足使用需求，或者即使单个服务器能够达到这个要求，但是其成本太高，所以往往使用虚拟服务器来平衡运用需求和成本。

多个虚拟服务器，可以通过虚拟机软件架设在同一个物理服务器上，而每个虚拟服务器又可以提供服务器的所有功能，这样每个虚拟服务器的成本相对更低。

运行时，由用户远程操作属于自己的那一块，而那一块对任何用户而言，就是一台“完整”的服务器，与真实独立的主机功能完全一样，用户一般通过空间域名网络申请虚拟服务器，只需对自己的信息进行远程维护，无需维护硬件、操作系统及通信线路。

## 第二节 信息采集设备

数字档案馆中的信息采集设备主要是传统载体档案数字化设备及条码扫描和射频感应设备等，传统载体档案数字化设备主要有扫仪器、视频采集设备和用于实物档案拍摄的数码相机。本书主要介绍在数字档案馆建设和运行中普遍使用的设备，如扫描仪、存储设备等。条码扫描及射频感应设备尚未大规模应用，数码相机原理简单，本书不做介绍。

### 一、扫描仪

#### 1. 扫描仪原理

扫描仪是桌面模拟/数字转换器，它能获取模拟视觉信号，即文件、杂志、

照片、图片等，并将它们转化为计算机可识别的数字化的图形图像。随它捆绑的软件或专业图像处理软件（如 PhotoShop、Photostyler 等）可帮助用户扫描获取数据、修改编辑扫描得到的数据，并把处理后的数据用在计算机的其他应用上（如传真、排版等）。甚至可利用光学字符识别软件（Optical Character Recognition, OCR）扫描文本，把图像化的文本转化为计算机上的一般文本数据，以此实现文本输入的“视觉化”。

扫描仪获取数据的方法：光源水平地排成一行，光线穿过扫描仪表面、照亮目标文件，反射光通过透镜，这里有一排电荷耦合器件（charge-coupled device, CCD）检测反射光，CCD 将反射光转化为电脉冲，电脉冲的电压随反射光强度变化而改变。每个像元都有一个 CCD，600dpi（每英寸有 600 个点，Dot Per Inch）的扫描仪有 600 个 CCD 排列在扫描仪的感光头上。最后，一个或多个模/数转换器（A/D）将模拟电信号转化为计算机所能处理、识别的 0、1 数字信号。

## 2. 扫描仪的性能

### （1）分辨率

分辨率是扫描仪最主要的技术指标，它表示扫描仪对图像细节上的表现能力，即决定了扫描仪所记录图像的细致度，其单位为 PPI（Pixels Per Inch，每英寸像素）。通常用每英寸长度上扫描图像所含有像素点的个数来表示。大多数扫描的分辨率在 300~2400PPI 之间。PPI 数值越大，扫描的分辨率越高，扫描图像的品质越高，但这是有限度的。当分辨率大于某一特定值时，只会使图像文件增大而不易处理，并不能对图像质量产生显著的改善。对于丝网印刷应用而言，扫描到 600PPI 就已经足够。

扫描分辨率一般有两种：真实分辨率（又称光学分辨率）和插值分辨率。

光学分辨率就是扫描仪的实际分辨率，它决定图像清晰度和锐利度的关键性能指标。

插值分辨率则是通过软件运算的方式来提高分辨率的数值，即用插值的方法将采样点周围遗失的信息填充进去，因此也称作软件增强的分辨率。例如，扫描仪的光学分辨率为 300PPI，则可以通过软件插值运算法将图像提高到 600PPI，插值分辨率所获得的细部资料要少些。尽管插值分辨率不如真实分辨率，但它却能大大降低扫描仪的价格，且对一些特定的工作，例如扫描黑白图

像或放大较小的原稿时十分有用。

400 PPI 的扫描分辨率指每英寸（1 英寸= 2.54cm）原稿被转换成 400 个像素。

通过简单的计算就可以算出多少尺寸原稿被转换成一个像素了，如下：

$$\frac{1\text{英寸}}{400\text{PPI}} = \frac{2.54\text{cm}}{400\text{PPI}} = 0.0000635\text{m/像素} = 63.5\mu\text{m/像素}$$

1 英寸或 2.54cm 被转换成 400 个像素。所以，1 米的 63,500,000 分之一被转换成一个像素。

如果 400PPI 的分辨率同时适用于主扫描方向和副扫描方向，通过扫描产生的一个像素将与一块 63.5μm×63.5μm 的原稿相对应。

## （2）灰度级

灰度级表示图像的亮度层次范围。扫描仪级数越多扫描图像的亮度范围越大、层次越丰富，多数扫描仪的灰度为 256 级。256 级灰阶中可以真实呈现出比肉眼所能辨识出来的层次更多的灰阶层次。

## （3）扫描仪色彩数

色彩数表示彩色扫描仪所能产生颜色的范围。通常用表示每个像素点颜色的数据位数即比特位（bit）表示。所谓 bit，是计算机最小的存贮单位，以 0 或 1 来表示比特位的值，越多的比特位数可以表现越复杂的图像信息。例如，常说的真彩色图像指的是每个像素点由三个 8 比特位的彩色通道所组成的 24 位二进制数表示，红绿蓝通道结合可以产生 2\*24=16.67M（兆）种颜色的组合，色彩数越多扫描图像越鲜艳真实。

## （4）扫描仪扫描速度

扫描速度有多种表示方法，因为扫描速度与分辨率、内存容量、缓存存取速度以及显示时间、图像大小等有关，通常用指定的分辨率和图像尺寸下的扫描时间来表示。

## （5）扫描仪扫描幅面

表示扫描图稿尺寸的大小，常见的有 A4、A3、A0 幅面等。

扫描仪与计算机连接的方式一般有三种，即并行端口接口、SCSI 接口、USB 接口。

# 3. 扫描仪的选购

扫描仪能将图片、文稿、照片、胶片、图纸等图形文件输入计算机，由于

普及型扫描仪已降至千元以下，使扫描仪逐步成为数字档案馆建设中不可缺少的设备。

#### （1）性能参数

描述扫描仪的性能参数很多，以下介绍一般用户购买时需要考虑的技术指标。

##### ① 扫描幅面

扫描幅面通常有 A4、A4 加长、A3、A1、A0 等规格。大幅面扫描仪价格很高，但在数字档案馆建设中，如果数字化工作不外包，A4 和 A3 幅面的扫描仪是必选的。至于是否采购更大幅面的扫描仪，应根据本单位需数字化档案的规格来确定。

##### ② 分辨率

分辨率反映扫描图像的清晰程度。分辨率越高的扫描仪，扫描出的图像越清晰。采购时，扫描仪的分辨率用每英寸长度上的点数 DPI (DotPerInch) 表示。一般办公用户建议选购分辨率为 600×1200（水平分辨率×垂直分辨率）的扫描仪。水平分辨率由扫描仪光学系统真实分辨率决定，垂直分辨率由扫描仪传动机构的精密程度决定，选购时主要考察水平分辨率。

##### ③ 色彩位数

色彩位数反映对扫描出图像色彩的区分能力。色彩位数越高的扫描仪，扫描出的图像色彩越丰富。色彩位数用二进制位数表示。例如，1 位的图像，每个像素点可以携带 1 位的二进制信息，只能产生黑或白两种色彩。8 位的图像可以给每个像素点 8 位的二进制信息，可以产生 256 种色彩。常见扫描仪色彩位数有 24 位、30 位、36 位和 42 位等常见标准。

##### ④ 感光元件

感光元件是扫描仪的眼睛，扫描质量与扫描仪采用的感光元件密切相关，普通扫描仪采用的感光元件有 CCD (ChargeCoupledDevice) 和 CIS (ContactImageSensor)。CCD 感光元件的扫描仪技术成熟。它配合由光源、几个反射镜和光学镜头组成的成像系统，在传感器表面进行成像，有一定景深，能扫描凹凸不平的实物。CIS 是广泛应用于传真机感光元件，其极限分辨率为 600DPI 左右，较 CCD 技术存在一定的差距，仅用于低档平板扫描仪。

#### （2）配套软件

扫描仪的功能都要通过相应的软件来实现，除驱动程序和扫描操作界面以外，几乎每一款扫描仪都会赠送一些图像编辑软件、OCR 等软件。不同扫描仪

配供软件性能、操作方法不一,对不熟悉图形处理的用户,建议选择配套操作简单,使用方便的扫描仪。

### (3) 接口

接口指扫描仪与电脑的连接方式,常见的有 SCSI 接口、EPP 接口和 USB 接口。

SCSI 接口扫描仪通过 SCSI 接口卡与电脑相连,数据传输速度快。缺点是安装较为复杂,需要占用一个扩展插槽和有限的电脑资源(中断号和地址)。如果经常扫描大量的图档,应当首选 SISC 接口扫描仪,可节约不少时间。

EPP 接口(打印机并口)用电缆即可连接扫描仪,安装简便;但其数据传输速度略慢于 SCSI 接口扫描仪。对扫描速度要求不高、扫描量不大、不经常使用扫描仪的用户,建议选购 EPP 接口扫描仪。

USB 接口扫描仪速度快、支持即插即用,与电脑的连接非常方便。

## 4. 扫描仪的分类

常见的扫描仪有平板扫描仪、高速扫描仪、工程扫描仪、书刊扫描仪、胶片扫描仪。

### (1) 平板扫描仪

平板扫描仪是目前最流行、使用最广的扫描仪。平板扫描仪通过移动感光头以扫过整张纸张来实现扫描。平板扫描仪的精确度是最高的,因为扫描时文件并不移动。另外,感光头在一个密封的环境中沿着一条粗细控制的路径移动,可提高实际分辨率。一般情况下,平板式扫描仪需要手动送纸。不过,现在也有可自动将纸送入送出到扫描仪面板的文件送纸器,实现平板扫描仪的自动送纸。在档案数字化处理中,平板扫描仪不损档案,可以适用不同幅面的档案扫描,应用比较广泛,是档案数字化必备的设备,但平板扫描仪速度较慢,一般用于扫描不能用高速扫描仪扫描的超过 A3 幅面的档案,或是珍贵的小幅面档案。

### (2) 高速扫描仪

高速扫描仪工作原理同平板扫描仪,但高速扫描仪可自动送纸,扫描时文件移动,但镜头不移动;由于采用自动送纸,扫描速度远高于平板扫描仪。现在大量 A3 幅面以下大多采用高速扫描仪进行数字化处理。

### (3) 工程扫描仪

工程扫描仪的工作原理与高速扫描仪基本相同,一般采用自动送纸,但扫

描文件幅面为 A0。扫描时文件移动，但镜头不移动。

#### （4）高拍仪

高拍仪扫描原理类似于数码相机，但镜头和置物台固定，相当于镜头和原稿台固定的数码相机。高拍仪扫描的特点是不需拆卷便可进行档案扫描。

#### （5）书刊扫描仪

书刊扫描仪工作原理类似于数码相机和高拍仪，但书刊扫描仪可能有多个镜头。书刊扫描仪必须是非接触式扫描方式，不需拆卷便可扫描，可以更好地保护档案。但书刊扫描仪需要手工翻页，速度慢于高速扫描仪，但快于平板扫描仪，见图 4-3。



图 4-3 某品牌书刊扫描仪

#### （6）胶片扫描仪

胶片扫描仪用于从摄影胶片或幻灯片上进行扫描，对图像进行处理以获得数字图像。胶片扫描仪的工作原理与平板扫描仪类似，胶片扫描仪与平板扫描仪最大的差别在于扫描材料的种类不同，前者是胶片，后者是诸如纸张等反射性媒质材料。若将胶片扫描仪的扫描头安装在平板扫描仪的顶部，则既能扫描需反射光源的纸质材料，又能扫描需直射光源的胶片。胶片扫描仪一般用于对缩微胶片数字化处理。

## 二、视频采集转换设备

视频采集转换设备主要是视频采集卡(Video Capture card),也称视频卡,将模拟摄像机、录像机、LD 视盘机、电视机输出的视频信号等输出的视频数据或视频和音频的混合数据输入电脑,并转换成电脑可辨别的数字数据,存储在电脑中,成为可编辑处理的视频数据文件。按照其用途可以分为广播级视频采集卡、专业级视频采集卡、民用级视频采集卡。

### 1. 视频采集卡用途

在电脑上通过视频采集卡可以接收来自视频输入端的模拟视频信号,对该信号进行采集、量化成数字信号,然后压缩编码成数字视频,具体原理见图 4-4。

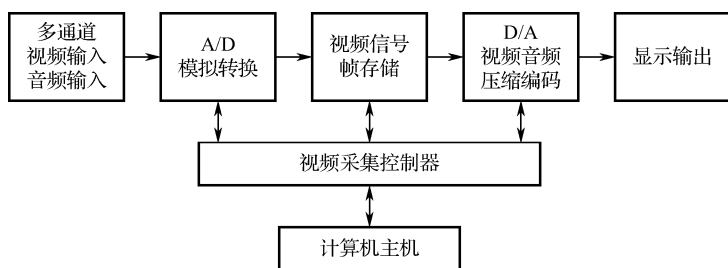


图 4-4 视频采集卡数模 A/D 转换流程

大多数视频卡都具备硬件压缩的功能,在采集视频信号时首先在卡上对视频信号进行压缩,然后再通过 PCI 接口把压缩的视频数据传送到主机上。一般的 PC 视频采集卡采用帧内压缩的算法把数字化的视频存储成 AVI 文件,高档一些的视频采集卡还能直接把采集到的数字视频数据实时压缩成 MPEG-1 格式的文件。

视频采集卡是进行视频处理必不可少的硬件设备,是视频数字化后期制作中必不可少的硬件设备。

通过视频采集卡,可以把摄像机拍摄的视频信号或录像机播放的视频信息转存到计算机中。利用相关的视频编辑软件,对数字化的视频信号进行后期编辑处理,例如,剪切画面,添加滤镜、字幕和音效,设置转场效果,以及加入



各种视频特效等，最后将编辑完成的视频信号转换成标准的 VCD、DVD 以及网上流媒体等格式，方便传播和保存。

## 2. 视频采集卡分类

视频采集卡按照视频信号源分，可以分为数字采集卡（使用数字接口）和模拟采集卡。

视频采集卡按照安装方式分，可以分为外置采集卡（盒）和内置式板卡。

视频采集卡按照视频压缩方式分，可以分为软压卡（消耗 CPU 资源）和硬压卡。

视频采集卡按照视频信号输入输出接口分，可以分为 1394 采集卡、USB 采集卡、HDMI 采集卡、VGA 视频采集卡、PCI 视频卡、PCI-E 视频采集卡。

视频采集卡按照其用途可分为广播级视频采集卡、专业级视频采集卡、民用级视频采集卡。它们的区别主要是采集的图像指标和采集图像的质量不同。

广播级视频采集卡的最高采集分辨率一般为 768X576（均方根值）PAL 制，或 720X576（CCIR-601 值）PAL 制 25 帧每秒，或 640X480/720X480 NTSC 制 30 帧每秒，最小压缩比一般在 4:1 以内。此类产品的特点是采集的图像分辨率高，视频信噪比高，缺点是视频文件大，每分钟数据量至少为 200MB。

专业级视频采集卡的级别比广播级视频采集卡的性能稍微低一些，两者的分辨率相同，但压缩比稍微大一些，其最小压缩比一般在 6:1 以内，输入输出接口为 AV 复合端子与 S 端子。

民用级视频采集卡的动态分辨率一般最大为 384X288，PAL 制 25 帧每秒。

## 3. 内、外置式视频采集卡

外置式的 USB 视频采集卡和内置式的 PCI 接口视频接收卡各有优缺点。首先，内置式视频接收卡价格比外置的便宜；其次，内置式视频采集卡不会占用外部的桌面空间；其三，内置视频卡不需外接电源；其四，内置的视频接收卡大多数都可以在 Windows 系统下实现多任务。但内置式视频卡也有其弱点。一是易受到电脑内部元器件的电磁干扰，引起其播放质量下降；二是在板卡安装时要比外置的麻烦一些，必须拆开机箱才可以进行安装，在安装驱动或软件时也容易和其他电脑设备发生冲突。外置式采集卡的优点就显而易见，外置式性能比起内置的视频接收卡来说会稳定一些。

#### 4. 视频采集卡系统配置

视频采集卡是视频采集和压缩同步进行，也就是说视频流在进入电脑的同时就被压缩成 MPG 格式文件，这个过程就要求电脑有高速的 CPU、足够大的内存、高速的硬盘、通畅的系统总线。

#### 5. 选购要点

##### （1）选择制式

视频采集卡根据其结构的不同可以分为内置和外置两种制式，外置式视频卡也叫视频接收盒，它是一个相对独立的设备，大都可以独立于电脑主机工作，无须打开计算机和运行软件就可以接受视频信息，在附加功能上都提供 AV 端子和 S 端子输入、多功能遥控、多路视频切换等。外置视频盒安装和操作简单，视频清晰度优于内置产品。内置的视频卡除提供标准视频接收功能外，往往提供不同程度的视频捕捉功能，可以把捕捉动态/静态的视频信号转换成数据流。具备视频捕捉的视频卡在接收视频信息之余，还能配合模拟制式摄像装置构成可视通信系统。

##### （2）选择捕捉效果

在捕捉效果上，应选择动态捕捉效果更接近于标称的 30 帧/秒或 25 帧/秒的产品为好，这其中还应保证捕捉到的图像画面的色彩、亮度、对比度的失真最小。

##### （3）选择分辨率

视频采集卡的分辨率还与所联的电脑密不可分，如果要通过视频采集卡来获取一些高质量的视频画面，应该留意视频采集卡在播放动态视频时的分辨率大小，分辨率越高的则越好，如果要实现高清效果，建议选用 hdmi 采集卡。最低要求不能低于普通电视的分辨率标准。VGA 高清采集卡可逐行采集 1920×1440×60HZ 的 VGA 信号，为使视频卡能出现最完美的演示效果，建议将计算机的分辨率调整到与其所购买的视频采集卡的分辨率一致。

##### （4）视频格式

如果购买的视频卡主要用于进行视频编辑处理，就要注意采集卡捕捉影像之后可以转存的视频格式，有的视频采集卡只能保存为 AVI 一种视频格式，并且没有影像压缩功能。为了能适应多种格式的视频信息的编辑处理，分辨率高，

可以保存为多种影像，有图像压缩功能的视频采集卡。

## 6. 主要问题

丢帧现象由于模拟视频输入端可以提供不间断的信息源，视频采集卡要采集模拟视频序列中的每帧图像，并在采集下一帧图像之前把这些数据传入计算机。因此，实现实时采集的关键是每一帧所需的处理时间。如果每帧视频图像的处理时间超过相邻两帧之间的相隔时间，则会出现数据丢失，也即丢帧现象。采集卡都是把获取的视频序列先进行压缩处理，然后再存入硬盘，也就是说视频序列的获取和压缩是在一起完成的，免除了再次进行压缩处理的不便。不同档次的采集卡具有不同质量的采集压缩性能。

# 第三节 存储设备

存储和还原输出设备也是数字档案馆运行必不可少的设备，由于还原输出设备比较简单，这里不做专门叙述，本节重点对存储设备进行叙述。存储设备按用途分类，可分为在线存储设备、近线存储设备、离线存储设备。

## 一、各类存储设备原理

### 1. 硬盘

硬盘是计算机主要的存储媒介之一，由一个或者多个铝制或者玻璃制的碟片组成。碟片外覆盖有铁磁性材料。

硬盘有固态硬盘（SSD，新式硬盘）、机械硬盘（HDD，传统硬盘）、混合硬盘（HHD，一块基于传统机械硬盘诞生出来的新式硬盘）。SSD 采用闪存颗粒来存储，HDD 采用磁性碟片来存储，混合硬盘（HHD，Hybrid Hard Disk）是把磁性硬盘和闪存集成到一起的一种硬盘。绝大多数硬盘都是固定硬盘，永久性地密封固定在硬盘驱动器中。硬盘选购一般考虑以下基本参数。

### （1）硬盘容量

作为计算机系统的数据存储器，容量是硬盘最主要的参数。

硬盘的容量以兆字节(MB/MiB)、千兆字节(GB/GiB)或百万兆字节(TB/TiB)为单位，而常见的换算式为：1TB=1024GB，1GB=1024MB 而 1MB=1024KB。但硬盘厂商通常使用的是 GB，且 1G=1000MB，而 Windows 系统仍以“GB”字样来表示“GiB”单位（1024 换算的），因此，在 BIOS 中或在格式化硬盘时看到的容量会比厂家的标称值小。

硬盘的容量指标还包括硬盘的单碟容量。所谓单碟容量是指硬盘单片盘片的容量，单碟容量越大，单位成本越低，平均访问时间也越短。

一般情况下硬盘容量越大，单位字节的价格就越便宜，但超出主流容量的硬盘会例外。

### （2）硬盘转速

转速（Rotational Speed 或 Spindle speed）是硬盘内电机主轴的旋转速度，也就是硬盘盘片在一分钟内所能完成的最大转数。转速的快慢是标示硬盘档次的重要参数之一，它是决定硬盘内部传输率的关键因素之一，在很大程度上直接影响硬盘的速度。硬盘的转速越快，寻找文件的速度也就越快，相对的硬盘的传输速度也就得到了提高。硬盘转速以每分钟多少转来表示，单位为 RPM，RPM 是 Revolutions Per minute 的缩写，是转/每分钟。RPM 值越大，内部传输速度就越快，访问时间就越短，硬盘的整体性能也就越好。

普通硬盘转速一般有 5400rpm、7200rpm，这几种高转速硬盘也是台式机用户的首选；而对于笔记本用户则是以 4200rpm、5400rpm 为主，虽然已经有公司发布了 10000rpm 的笔记本硬盘，但在市场中还较为少见；服务器用户对硬盘性能要求最高，服务器中使用的 SCSI 硬盘转速基本都采用 10000rpm，甚至还有 15000rpm 的，性能要超出家用产品很多。较高的转速可缩短硬盘的平均寻道时间和实际读写时间，但随着硬盘转速的不断提高，也带来了温度升高、电机主轴磨损加大、工作噪音增大等负面影响。

### （3）硬盘平均访问时间

平均访问时间（Average Access Time）是指磁头从起始位置到到达目标磁道位置，并且从目标磁道上找到要读写的数据扇区所需的时间。

平均访问时间体现了硬盘的读写速度，它包括了硬盘的寻道时间和等待时间，即：平均访问时间=平均寻道时间+平均等待时间。

硬盘的平均寻道时间（Average Seek Time）是指硬盘的磁头移动到盘面指

定磁道所需的时间。这个时间越小越好，硬盘的平均寻道时间通常在 8ms 到 12ms 之间，而 SCSI 硬盘则应小于或等于 8ms。

硬盘的等待时间，又叫潜伏期（Latency），是指磁头已处于要访问的磁道，等待所要访问的扇区旋转至磁头下方的时间。平均等待时间为盘片旋转一周所需的时间的一半，一般应在 4ms 以下。

#### （4）硬盘传输速率

传输速率（Data Transfer Rate）是指硬盘读写数据的速度，单位为兆字节每秒（MB/s）。硬盘数据传输率包括内部数据传输率和外部数据传输率。

内部传输率（Internal Transfer Rate）也称持续传输率（Sustained Transfer Rate），它反映了硬盘缓冲区未用时的性能，内部传输率主要依赖于硬盘的旋转速度。

外部传输率（External Transfer Rate）也称突发数据传输率（Burst Data Transfer Rate）或接口传输率，它标称的是系统总线与硬盘缓冲区之间的数据传输率，外部数据传输率与硬盘接口类型和硬盘缓存的大小有关。

Fast ATA 接口硬盘的最大外部传输率为 16.6MB/s，而 Ultra ATA 接口的硬盘则达到 33.3MB/s。

#### （5）硬盘缓存

缓存（Cache memory）是硬盘控制器上的一块内存芯片，具有极快的存取速度，它是硬盘内部存储和外界接口之间的缓冲器。由于硬盘的内部数据传输速度和外界介面传输速度不同，缓存在其中起到一个缓冲的作用。缓存的大小与速度是直接关系到硬盘的传输速度的重要因素，能够大幅度地提高硬盘整体性能。当硬盘存取零碎数据时需要不断地在硬盘与内存之间交换数据，有大缓存，则可以将那些零碎数据暂存在缓存中，减小系统的负荷，也提高了数据的传输速度。

#### （6）硬盘存储电子档案的特性

硬盘一般用于电子档案在线存储，但近来也有厂商将硬盘用电子档案的离线和近线存储。有厂商开发了用于近线存储的磁盘库，正在档案部门试用。

## 2. 磁带

以磁带为存储介质，由磁带机及其控制器组成的存储设备，是计算机的一种辅助存储器。磁带机由磁带传动机构和磁头等组成，能驱动磁带相对磁头运

动，用磁头进行电磁转换，在磁带上顺序地记录或读出数据。磁带存储器是计算机外围设备之一，以顺序方式存取数据。磁带控制器是中央处理器在磁带上存取数据用的电路控制装置。存储数据的磁带可脱机保存和互换读出。

一盘磁带有始端标记（BOT）和尾端标记（EOT），中间可记若干个文件。每个文件由 1 至若干个数据块组成，两个文件之间有带标隔开。在数据块之间和数据块与带标之间有规定长度的空白间隔。数据块内数据有规定的格式。数据块是磁带读写数据的基本单元。磁头在空白间隔内启停。通常带宽为 12.7 毫米，有 9 个磁道，8 道记数据，1 道记数据的奇偶位，数据以字节串联记入。常用的标准格式有不归零（NRZI）、相位编码（PE）和成组编码（GCR）三种。记录密度分别为每毫米 32 位、63 位和 246 位，它们各自有规定的检错和纠错方法。

磁带库是像自动加载磁带机一样的基于磁带的备份系统，磁带库由多个驱动器、多个槽、机械手臂组成，并可由机械手臂自动实现磁带的拆卸和装填。它能够提供同样的基本自动备份和数据恢复功能，但同时具有更先进的技术特点。它可以多个驱动器并行工作，也可以几个驱动器指向不同的服务器来做备份，存储容量达到 PB（1PB=100 万 GB）级，可实现连续备份、自动搜索磁带等功能，并可在管理软件的支持下实现智能恢复、实时监控和统计，是集中式网络数据备份的主要设备。

磁带库不仅数据存储量大，在备份效率和人工占用方面同样拥有优势。在网络系统中，磁带库通过 SAN（Storage Area Network，存储局域网）系统可形成网络存储系统，很容易完成远程数据访问、数据存储备份，或通过磁带镜像技术实现多磁带库备份，无疑是数据仓库、ERP 等大型网络应用的良好存储设备。且磁带介质保存时间久远成本低廉，已广泛应用于银行、广播电视媒体、档案馆、国土资源、卫星资源等行业。

虚拟磁带库的概念其实非常简单，它在本质上是磁盘阵列硬件设备，但在软件功能上却模拟磁带备份的形式。因此，对于存储管理员来说，它就是一个磁带库，对它的管理与管理一个物理磁带库完全一样。

虚拟磁带库主要有以下优势。

（1）相对于机械磁带，虚拟磁带库使用了兼容磁带备份管理软件以及传统备份流程，这使得设备的可用性以及备份的可靠性得到了大幅提升。

（2）性能大幅提高，可支持接近磁盘阵列极限速度的备份以及恢复速度。而且免疫病毒，数据安全性等同普通磁带库。

(3) 恢复工作极为简便, 如果所需数据存在 VTL 当中, 则不会涉及任何机械工作, 恢复工作就像磁盘备份的速度一样。

(4) 虚拟磁带库采用基于 RAID 保护的磁盘阵列, 从而将备份的可靠性较常规磁带备份提高了若干量级。封闭式结构的磁盘介质本身的 MTBF (平均无故障间隔) 一般为开放式结构磁带介质的 5 倍以上。

(5) 兼容性好, 标准的 FC、SCSI 或 iSCSI 接口设备, 兼容流行的主机设备和操作系统。

(6) 实用性好, 与现有磁带库应用方式一致, 不用更改现有存储应用软件的管理策略。

(7) 虚拟磁带库用电子化的“机械手”和“磁带驱动器”, 代替了机械磁带库中裸露、易损的机械装置, 基于 RAID 保护的磁盘阵列具备降级工作能力, 且具自动报警和在线热恢复能力。

从以上虚拟磁带库的优势可以看出, 虚拟磁带库不仅解决了传统磁带库维护负担高、备份失效率高, 以及备份恢复能力不佳的问题, 而且也改变了磁盘备份容易被误删除或被病毒感染以及不便于在 SAN 环境中统一管理和优化使用的劣势。

### 3. U 盘

相较于其他可携式存储设备, U 盘有许多优点。占空间小, 操作速度较快 (USB1.1、2.0、3.0、3.1 标准), 能存储较多数据, 并且性能较可靠 (由于没有机械设备), 在读写时断开而不会损坏硬件, 只会丢失数据。

与其他的闪存设备相同, 闪存盘在总读取与写入次数上也有限制。中档的闪存盘在正常使用状况下可以读取与写入数十万次, 但当闪存盘变旧时, 写入的动作会更耗费时间。当用闪存盘来运行应用程序或操作系统时, 便应该考虑这点。有些程序开发者特别针对这个特性以及容量的限制, 为闪存盘撰写了特别版本的操作系统 (例如 Linux) 或是应用程序 (例如 Mozilla Firefox)。它们通常对使用空间做优化, 并且将暂存数据存储电脑的主存中, 而不是闪存盘里。

### 4. 光盘

光盘是以光信息作为存储载体并用来存储数据的一种载体。分不可擦写光

盘，如 CD-ROM、DVD-ROM 等；和可擦写光盘，如 CD-RW、DVD-RAM 等。

光盘结构分为基板、记录层、反射层、保护层。

基板是各功能结构（如沟槽等）的载体，其使用的材料是聚碳酸酯（PC），冲击韧性极好、使用温度范围大、尺寸稳定性好、耐候性、无毒性。一般来说，基板是无色透明的聚碳酸酯板，在整个光盘中，它不仅是沟槽等的载体，更是整个光盘的物理外壳。CD 光盘的基板厚度为 1.2mm、直径为 120mm，中间有孔，呈圆形，它是光盘的外形体现。光盘之所以能够随意取放，主要取决于基板的硬度。

记录层是烧录时刻录信号的地方，其主要的工作原理是在基板上涂抹专用的有机染料，以供激光记录信息。由于烧录前后的反射率不同，经由激光读取不同长度的信号时，通过反射率的变化形成 0 与 1 信号，借以读取信息。

一次性记录的 CD-R 光盘在进行烧录时，激光就会对在基板上涂的有机染料，进行烧录，直接烧录成一个接一个的“坑”，这样有“坑”和没有“坑”的状态就形成了 0 和 1 的信号，这一个接一个的“坑”是不能恢复的，也就是当烧成“坑”之后，将永久性地保持现状，这也就意味着此光盘不能重复擦写。这一连串的 0、1 信息，就组成了二进制代码，从而表示特定的数据。

在这里，需要特别说明的是，对于可重复擦写的 CD-RW，当激光在烧录时，就不是烧成一个接一个的“坑”，而是改变碳性物质的极性，通过改变碳性物质的极性，来形成特定的 0、1 代码序列。这种碳性物质的极性是可以重复改变的，这也就表示此光盘可以重复擦写。

反射层是光盘的第三层，它是反射光驱激光光束的区域，借反射的激光光束读取光盘片中的资料。其材料是纯度为 99.99% 的纯银金属。

保护层是用来保护光盘中的反射层及染料层防止信号被破坏。材料为光固化丙烯酸类物质。

印刷层是印刷盘片的客户标识、容量等相关资讯的地方，这就是光盘的背面，它不仅可以标明信息，还可以起到一定保护光盘的作用。

根据光盘结构，光盘主要分为 CD、DVD、蓝光光盘等几种类型，这几种类型的光盘，在结构上有所区别，但在原理是一致的。只读的 CD 光盘和可记录的 CD 光盘在结构上没有区别，它们主要区别在材料的应用和某些制造工序。

从主要结构来讲，CD、DVD 光盘的结构是一致的，只不过，它们的厚度和用料有所不同。

在实际应用中，读取和烧录 CD、DVD、蓝光光盘的激光是不同的。CD 的



容量只有 700MB 左右，而 DVD 则可以达到 4.7GB，而蓝光光盘更是可以达到 25GB。它们之间的容量差别，同其相关的激光光束的波长密切相关。

光盘片的记录密度受限于读出的光点大小，即光学的绕射极限（Diffraction Limit），其中包括激光波长，物镜的数值孔径。

对于 CD 光盘，其激光波长为 780nm，物镜的数值孔径为 0.45，激光束会集到一点的距离需要 1.2mm，这就决定了 CD 光盘基板的厚度为 1.2mm。

DVD 光盘的激光波长为 650nm，物镜的数值孔径为 0.6，而激光束会集到一点的距离只需要 0.6mm，这决定 DVD 光盘基板的厚度为 0.6mm。不过，0.6mm 的厚度太薄，其制造出来的光盘也会因为太薄而容易折断。因此，在 DVD 的实际制造过程中，会把两片 0.6mm 厚的基板叠合在一起，共同组成 1.2mm 的厚度，一片基板记录数据，另一片基板完全起保护作用。

光盘的发展趋势是向大容量存储发展，业界的技术研发也以此为导向。

已经出现了单面双层的 DVD 盘片。单面双层盘片（DVD+R Double Layer）是利用激光（Laser beam）聚焦的位置不同，在同一面上制作两层记录层，单面双层盘片在第一层及第二层的激光功率（Writing Power）相同（激光功率为 <30mW），反射率（Reflectivity）也相同（反射率为 18%~30%），刻录时，可从第一层连续刻录到第二层，实现信息刻录不间断。

## 二、在线存储

在线存储的优点是近零访问延迟，支持大量用户访问，容易集中管理，电子档案管理系统日常运行所用的存储，均是在线存储。

### 1. DAS

DAS 是 Direct Attached Storage 的缩写，即“直接连接存储”，是指将外置存储设备通过连接电缆，直接连接到一台计算机上。采用直接外挂存储方案的服务器结构如同 PC 机架构，外部数据存储设备采用 SCSI 技术，或者 FC 技术，直接挂接在内部总线上的方式，数据存储是整个服务器结构的一部分，数据和操作系统都未分离。DAS 这种直连方式，能够解决单台服务器的存储空间扩展、高性能传输需求。随着大容量硬盘的推出，单台外置存储系统容量还会上升。此外，DAS 还可以构成基于磁盘阵列的双机高可用系统，满足数据存储对高可

用的要求。从趋势上看，DAS 仍然会作为一种存储模式，继续得到应用。

## 2. NAS

NAS 是英文 Network Attached Storage 的缩写，通常译为网络附加存储。NAS 作为一种概念，是 1996 年从美国硅谷提出的，其主要原理是把存储设备和网络接口集成在一起，直接通过以太网网络存取数据。即把存储功能从通用文件服务器中分离出来，使其更加专门化，从而获得更高的存取效率，更低的存储成本。

NAS 设备有可靠稳定的性能，特别优化的文件管理系统和低廉的价格，它采用的嵌入式技术，使得 NAS 具有性能稳定、功能专一的特点。

NAS 设备内置优化的独立存储操作系统，可以有效、紧密地释放系统总线资源，全力支持 I/O 存储，同时 NAS 设备一般集成本地的备份软件，可以不经过服务器将 NAS 设备中的重要数据进行本地备份，而且 NAS 设备提供硬盘 RAID、冗余的电源和风扇以及冗余的控制器，可以保证 NAS 的稳定应用。

NAS 设备主要用来实现在不同操作系统平台下的文件共享应用，与其他服务器或 DAS 存储设备相比，NAS 设备的安装、调试、使用和管理非常简单，采用 NAS 可以节省一定的设备管理与维护费用。NAS 设备提供 RJ-45 接口和单独的 IP 地址，可以将其直接挂接在主干网的交换机或其他局域网的 Hub 上，通过简单的设置（如设置机器的 IP 地址等）就可以在网络即插即用使用 NAS 设备，进行网络数据在线扩容时也无需中断，从而保证数据流畅存储。

NAS 在局域网以文件为单元进行数据存取，也就是说利用网络文件系统、TCP/IP 协议、以太网络设施，实现数据存取。它可以充分利用用户现有的局域网络设施，大大节省了用户在存储上的投资。但 NAS 也有弱点，即受网络带宽的限制，且 TCP/IP 的帧传输方式，使得带宽利用率不高。一般情况下，NAS 设备的数据传输带宽仅能达到 9~15MB/s，在 TCP/IP 技术上，以文件为单元进行传输，传输时的丢包现象限制了 NAS 的速度，甚至威胁到数据唯一性和安全。目前大部分的数据都是基于关系型数据库进行存储的，需要实时高速的数据读取和存储，一般数据库都采用“块”（Block）的方式进行数据传输，这样 NAS 就无法在主流数据库应用上。

### 3. SAN

SAN 是 Storage Area Network 的缩写，即“存储区域网络”。主机和存储装置之间通过智能型的光纤通道交换器连接，并与存储网络的管理软件统一管理，这种方式就是 SAN。它最大特性是将网络和设备的通信协议与传输物理介质隔离开。这样多种协议可在同一个物理连接上同时传送，高性能存储体和宽带网络使用单 I/O 接口，使得系统的成本和复杂程度大大降低。

因为采用了全双工串行通信原理传输数据 FC 技术，SAN 具有更高的带宽，在 1GB 标准下，传输速率高达 1062.5Mbps，即为 100MB/S，双环可达 200MB/S，2GB，上述数字将翻倍。FC 标准下可以通过同轴线、光纤介质进行设备间的信号传输，使用同轴线传输距离为 30 米，使用单模光纤传输距离可达 10km 以上，这使得在 SAN 模式下实现物理上分离的、不在机房的存储变得非常容易。

SAN 主要用于存储量大的工作环境，如电信、银行、电子政务的信息中心等。随着各种用户数据量的剧增，对存储在可用性、可扩展性、管理性上要求的提高。

## 三、近线存储

近线存储（NearStore），是随客户存储环境的细化所提出的一个概念，所谓的近线存储，外延相对较广泛，主要定位于客户在线存储和离线存储之间的应用。就是指将那些并不是经常用到，或者数据的访问量并不大的数据存放在性能较低的存储设备上。但同时对这些的设备要求是寻址迅速、传输率高。例如，一些长期保存的不常用的客户文件的归档。因此，近线存储对性能要求相对来说并不高，但又要求相对较好的访问性能。同时多数情况下由于不常用的数据要占总数据量比较大的比重，这就要求近线存储设备在需要容量相对较大。在业界传统定义的近线存储设备主要为 DVD-RAM 光盘塔和光盘库设备。但随着存储设备的不断发展，根据客户存储需求的不同，也会把低端的磁盘阵列（例如 DS4100）或高端的磁带设备（例如 3592）作为近线存储应用设备。

近线存储提供便宜、可靠、无限制的数据备份，并能减少在线存储访问量。对于个人和小型商户来说，如果用户可以忍受一些存储或者检索数据时产生的时间延迟，这将是一个理想的解决方案。近线存储媒体在架的时候是不会被病

毒感染的，因为这个时候的媒体在物理上是不连接在网络、计算机、服务器和因特网上的。当近线存储介质拿来覆盖设备的时候，它可以通过写保护来防止感染。如果用一台被感染的计算机来往近线存储介质上写入数据的话介质很可能会被感染。所以，推荐在使用前用杀毒软件扫描一下近线存储介质。

近线存储在数字档案馆系统中运用较多。它既避免了离线存储的慢速度，又避免了在线存储高价格，是数字档案馆建设中需要考虑的存储方案。当前近线存储的方案主要有光盘库和硬盘库。

## 四、离线存储

离线存储（OffStore）是用于对在线存储的数据进行备份，以防范可能发生的数据灾难，因此又称备份级的存储。电子档案离线存储一般采用光盘、硬盘、磁带。三种存储用于电子档案离线存储的优缺点比较，见表 4-2。

表 4-2 离线存储载体比较表

介 质	磁 带	光 盘	硬 盘
物理优点	易生产，使用广泛	数据不可修改，起到证据作用	存储容量大， 可长期保存
物理缺点	数据易受外界环境影响，保存时磁塑介质易粘连	片基易老化，盘片易划损	抗冲击力较弱，内部含有皮带、电容等器件需要定期使用
使用管理优点	系统成熟，存储容量大	通用性较高	信息易定位，定期备份 信息进快速，通用性较高
使用管理缺点	速度慢，定位信息困难，需专用设备读取，型号规格多样	使用中光盘数量大，读写速度较低	需定期启动保证正常工作
后期管理	需短周期对介质检验、读取、重新拷贝	需要对大量同类介质进行分类检验，并重新备份	通过技术对数据进行自检，重新拷贝数据速度也较快
保存	为保证数据的安全完整，需要的保存环境较苛刻	对保存环境要求较弱，但是长时间保存光盘基片易老化，导致盘片损坏	较易保存，且保存时占用空间较小
价格	需专门磁带设备读取，设备较贵，磁带价格在 650 元/盘左右	使用普遍，一张 DVD 盘容量为 4.7G，价格在 15 元/张左右	使用普遍，希捷 1T 普通型硬盘价格为 300 元/块左右

## 五、云存储

随着信息技术的发展，云技术逐渐应用到电子文件管理领域，为电子文件管理提供基础设施、硬件、软件以及平台等多方面的支持。但是，在云环境中如何管理电子文件提出了管理与技术的双重挑战，云服务如何保障电子文件的可信、真实、完整、可靠、有效与安全，成为采用云服务时重点考虑的问题，尤其是其在安全与保密方面的风险，使用户采用云服务时存在顾虑。

基于云在数字档案馆中应用的详细论述，请参考本书第十一章有关内容。

# 第五章 传统载体档案数字化

## 第一节 概述

传统载体档案数字化是数字档案馆建设的又一重点工作，是工作量最大的一项工作，也是使数字档案馆发挥成效地关键性工作。其主要任务是将传统载体的档案信息转换成数字信息并数据库化。传统载体档案数字化包括档案目录的数字化和档案全文的数字化。按载体种类分，传统档案数字化包括纸质档案数字化、多媒体（主要包括录音、录像）档案数字化和实物档案数字化。

### 一、档案数字化原则

早期，档案工作信息化曾经走过一段不小的弯路，把大量的精力投入到传统载体档案数字化工作当中，对馆（室）藏所有档案进行数字化，最后投入与产出不成比例，造成人力、物力的浪费。经过十几年的发展，档案界逐渐总结出若干传统载体档案数字化的原则。

#### 1. 规范化原则

规范化原则是指档案数字化时数据的格式、信息代码和元数据项必须规范一致，以方便数据管理、检索利用和数据交换与共享。相反，数据格式、信息代码和元数据项不统一、不规范，将不利于档案的管理、检索利用和数据交换共享。

统一的规范包含三个方面，一是数字化的档案信息的数据格式、信息代码和元数据项必须统一；二是数字化中应尽可能采用国家或国际标准，将数据格式、信息代码和元数据统一到国家或国际标准上来；三是数字化过程要规范，每一个环节都必须遵循相应的规范。

## 2. 准确性原则

档案著录的信息应与档案原有属性信息相一致，建立的对应关系应准确。

## 3. 完整性与系统性原则

档案数字化过程中录入信息的完整性，确保内容信息、背景信息和元数据等信息的系统性，能够全面地记录和描述档案的各种属性。

## 4. 需求牵引，成本效益原则

档案数字化以档案利用为目的，贯彻需求牵引和成本效益原则，根据需要，先数字化利用率高的档案，后数字化利用率低的档案。根据这一原则，应贯彻“五个优先”原则，下列档案应优先考虑建立数据库。

### （1）常用优先

这是贯彻需求牵引原则，把利用频率较高的档案优先数字化，以满足当前利用的需求，增加数字档案馆的效益。

### （2）目录优先

即把档案目录优先数字化，以满足检索需求。目录优先还体现为成本效益的需要。由于全文数字化建设需投入的人力物力较大，见效比较慢，如优先将档案的目录进行数字化，实现档案的计算机检索，以尽快见到效果。

### （3）急用优先

企业在某一个时期有特定的专项任务，从而派生出对某一方面档案利用的紧迫需求，档案数字化时应考虑这一需求，将急用的档案优先数字化，以满足紧急需求。如产品开发时期有可能对某一老产品的档案调用较多，可考虑优先将该老产品档案数字化，以方便利用和满足业务发展对档案利用的需求。

### （4）孤本优先

即把珍贵的档案优先数字化，以满足珍藏需求。数字化后，孤本档案利用可不用调用原件，可有效保护档案原件，达到珍藏的目的。

### （5）特色优先

即将体现本企业特色的档案优先数字化，以满足特殊需求。不少企业档案机构保存有称之为“镇馆之宝”的档案，这部分档案较易引起利用者的兴趣，可优先数字化，有利于增强档案部门的影响力。

## 5. 安全保密原则

传统载体数字化前，应建立安全保障和保密措施，以确保数字化过程中的档案信息的保密。安全保密包括安全与保密两个方面，其含义包括以下几个方面。

(1) 要建立在档案数字化过程中的档案信息安全保密体系。

建立一套管理制度与机制，从人、物、技术、管理等方面保证档案信息安全、防止信息失泄密。

(2) 要注重档案数字化过程中档案信息的安全。

防止传统载体档案及数字化后信息丢失、损毁。保证传统载体安全，注意数字化场所的安全，注意机器、人员对档案造成的损毁；建立档案出入、交接登记制，分清档案安全保管责任。数字化后，需做好信息备份工作，及时备份数据；防止计算机病毒破坏档案数据信息。

(3) 要注重档案数字化过程中档案信息的保密。

要从人防、物防、技防和加强管理入手，在人防方面要为档案数字化配备素质可靠的工作人员；在物防方面要选择具有保密措施的环境作为档案数字化建设的场所，场所要符合保密要求；在技防方面要特别注意用于档案数字化信息系统的保密建设，投入使用的电子档案管理系统要符合涉密管理要求。

## 6. 档案数字化“两个同步”原则

(1) 档案数字化转换与档案鉴定同步推进

在档案数字化过程中，要对每个保管单位进行梳理，在这个过程中对需要数字化的档案进行一次鉴定是十分必要的。通过鉴定重新审视其价值，尤其是已到期或将到保管期限的档案，通过鉴定可剔除价值较低的档案，将其排除在数字化范围之外。数字化过程中的鉴定也可列入日常管理的定期鉴定，可谓一举两得。

(2) 数字化转换与档案利用同步推进

由于档案数量庞大，档案数字化工作周期一般较长，如果等到全部档案数字化完成后将档案数据开放检索，数字档案馆的效益将长期得不到体现，增加缩减投资和项目失败的风险。为此，必须尽早将档案数据开放检索，实行数字化与档案利用同步进行，完成一批投入，开放一批，尽早使数字档案馆建设工



作见到成效。

二、档案数字化的组织

档案数字化是数字档案馆建设中工作量较大的一项工作，需要较多的人员和机器来完成，如果自主完成，应组织好扫描阶段参与工作的人员和设备。

1. 正确选择数字化设备

数字化设备的选择，一般从档案幅面、档案纸张质量、档案珍贵程度等几个因素来考虑。一般来说，A4 以下幅面、纸张质量较好的档案选用高速扫描仪；A4 以下幅面，但纸张质量不好，或是珍贵的档案，一般选用平板扫描。现在由于扫描仪性能提高，出现了 A3 幅面高速扫描仪，A3 幅面的档案也可以用高速扫描仪。大于 A3 幅面的档案一般用平板扫描仪扫描。大于 A0 幅面的档案一般选用工程扫描仪扫描。

如果不想拆分档案，可选用书刊描扫仪或高拍仪进行扫描。

扫描仪选用如图 6-1 所示。

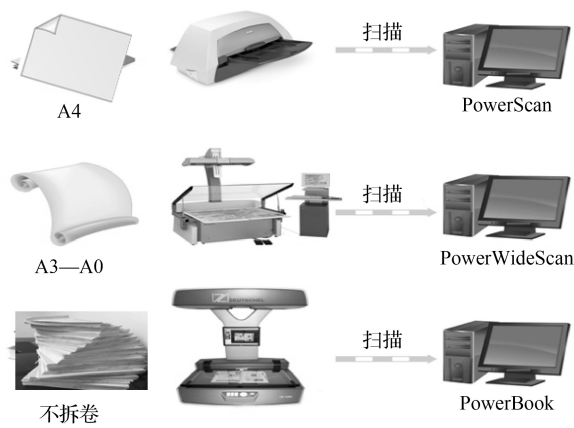


图 6-1 扫描仪选用图

2. 正确分配人员与数字化设备

在多人多设备情况下，采用协作方式开展扫描工作会大幅度提高工作效率，

因此，需要根据档案扫描流程，再结合设备进行合理的组织和分工。

#### （1）按幅面和工作内容进行分工

将人员分为档案拆装人员、目录录入人员、高速扫描人员、平板扫描人员和工程扫描人员；将需扫描的档案拆分为用高速扫描仪、平板扫描仪、工程扫描仪等种类，并分别由相应的人员承担，其中目录录入由专人承担，扫描完成后由档案拆装人员重新装订。这种按幅面结合流程分工组织扫描工作的方法效率高。但缺点是需要将成卷（盒）的档案分开，需要对分开的档案进行登记，恢复档案原状时易丢失，对档案安全不利，具体见图 6-2。

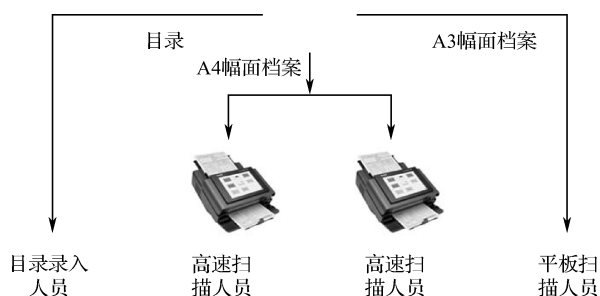


图 6-2 数字化处理分工图（按幅面）

#### （2）按保管单位分工

即不拆分档案，一卷（盒）档案的各种幅面档案的扫描、目录录入等工作全部由一个人承担。设备分派时，每人一台高速扫描仪，再设置一台或若干台公用的平板或工程扫描仪。每个人的工作流程是：先拆除装订物，然后目录录入、高速扫描，如遇不适合高速扫描的档案，则临时使用放置在附近的公用的平板或工程扫描仪处理。这样分工组织的好处，是不用将不同幅面档案分工，一个保管单位的档案一直在一个人手中，不流转，有利于档案的安全，工作质量、档案安全责任清晰，具体见图 6-3。

## 三、数字化技术

### 1. 扫描技术

扫描是用数字方式表现现实世界中的物体。大多数扫描设备用来扫描一张纸的表面或书本的一页，将纸张表面上的信息数字化。将这些信息数字化，就

是要产生与我们在一张纸上所看到的图像相对应的数字图像。数字图像只不过是描述一幅图像并可以用计算机存储和处理的一组二进制代码，它可以显示在计算机屏幕上，也可以用打印机打印出来。

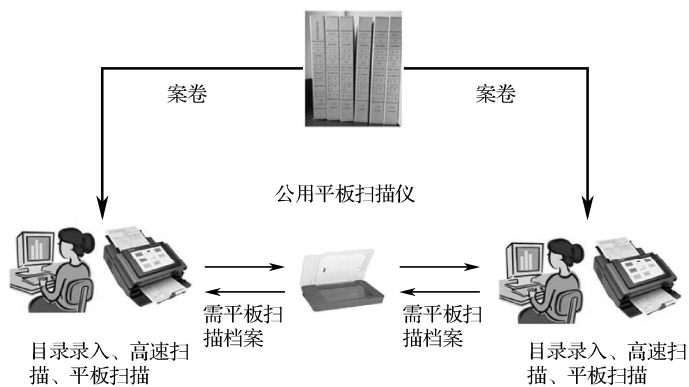


图 6-3 数字化处理分工图（按卷）

除了上面介绍的扫描仪外，还有用来扫描胶片和照片的扫描器，甚至还有扫描三维物体的扫描器。胶片扫描类似于扫描一张纸，其结果是数字图像；而三维扫描器则可以产生扫描物体的数字三维模型。

2. 音视频转换技术

电脑通过音视频采集卡可以接收来自视频输入端的模拟音视频信号，对该信号进行采集、量化成数字信号，然后压缩编码成数字音视频，这就是音视频转换技术。多媒体档案数字化就是采用音视频转换技术，将模拟信号的音视频转成数字格式，存储于电子档案管理系统中。

四、扫描文件存储格式

1. RAW 格式

RAW 是一种无损压缩格式，它的数据是没有经过处理的原文件，因此，它的大小要比 TIFF 格式略小。所以，当上传到电脑之后，要用图像软件的 Twain 界面直接转换成 TIFF 格式才能处理。

## 2. 图像格式 BMP 格式

BMP (BitMaP, 位图格式) 是一种与硬件设备无关的图像文件格式, 使用非常广泛。它采用位映射存储格式, 除了图像深度可选以外, 不采用其他任何压缩, 因此, BMP 文件所占用的空间很大。BMP 文件的图像深度可选 1bit、4bit、8bit 和 24bit。BMP 文件存储数据时, 图像的扫描方式是按从左到右、从下到上的顺序。

由于 BMP 文件格式是 Windows 环境中交换图像有关数据的一种标准, 因此, 在 Windows 环境中运行的图形图像软件都支持 BMP 图像格式。

## 3. PCX 格式

PCX (Personal Computer eXchange, 个人电脑交换格式) 图像文件的形成有一个发展过程。最先的 PCX 雏形是出现在 ZSOFT 公司推出的、名为 PC PAINBRUSH 的、用于绘画的商业软件包中, 后来微软公司将其移植到 Windows 环境中, 成为 Windows 系统的一个子功能。先在微软的 Windows3.1 中广泛应用, 随着 Windows 的流行、升级, 加之其强大的图像处理能力, 使 PCX 同 GIF、TIFF、BMP 图像文件格式一起, 被越来越多的图形图像软件工具所支持, 也越来越得到人们的重视。

PCX 是最早支持彩色图像的一种文件格式, 现在最高可以支持 256 种彩色, 显示 256 色的彩色图像。PCX 设计者超前地引入了彩色图像文件格式, 使之成为现在非常流行的图像文件格式。

## 4. TIFF 格式

TIFF (TagImage FileFormat, 标签图像文件格式), 是由 Aldus 和 Microsoft 公司为桌面出版系统研制开发的一种较为通用的图像文件格式。TIFF 格式灵活易变, 它又定义了四类不同的格式: TIFF-B 适用于二值图像; TIFF-G 适用于黑白灰度图像; TIFF-P 适用于带调色板的彩色图像; TIFF-R 适用于 RGB 真彩图像。

TIFF 支持多种编码方法, 其中包括 RGB 无压缩、RLE 压缩和 JPEG 压缩等。

TIFF 是现有图像文件格式中最复杂的一种, 它具有扩展性、方便性、可改性, 可以提供给 IBMPC 等环境中使用。

TIFF 图像文件由三个数据结构组成, 分别为文件头、一个或多个称为 IFD

的包含标记指针的目录，以及数据本身。

TIFF 图像文件中的第一个数据结构称为图像文件头或 IFH。这个结构是一个 TIFF 文件中唯一的、有固定位置的部分；IFD 图像文件目录是一个字节长度可变的信息块，Tag 标记是 TIFF 文件的核心部分，在图像文件目录中定义了所有图像参数，目录中的每一条目就包含图像的一个参数。

## 5. GIF 格式

GIF (Graphics Interchange Format, 图形交换格式), 是 CompuServe 公司在 1987 年开发的图像文件格式。GIF 文件的数据, 是一种基于 LZW 算法的连续色调的无损压缩格式。其压缩率一般在 50% 左右, 它不属于任何应用程序。目前几乎所有相关软件都支持它, 公共领域有大量的软件在使用 GIF 图像文件。

GIF 图像文件的数据是经过压缩的, 而且是采用了可变长度等压缩算法。所以, GIF 的图像深度从 1bit 到 8bit, 最多支持 256 种色彩的图像。GIF 格式的另一个特点, 是其在在一个 GIF 文件中可以保存多幅彩色图像, 如果把存于一个文件中的多幅图像数据逐幅读出并显示到屏幕上, 就可构成一种最简单的动画。

GIF 解码较快, 因为采用隔行存放的 GIF 图像, 在边解码边显示的时候可以分成四遍扫描。第一遍扫描虽然只显示了整个图像的八分之一, 第二遍的扫描后也只显示 1/4, 但这已经把整幅图像的概貌显示出来了。在显示 GIF 图像时, 隔行存放图像会感觉显示速度似乎要比其他图像快一些, 这是隔行存放的优点。

GIF 格式的文件是 8 位图像文件, 最多为 256 色, 不支持 Alpha 通道, 产生的文件较小, 常用于网络传输, 在网页上见到的图片大多是 GIF 和 JPEG 格式。与 JPEG 格式相比, GIF 格式的优点在于其文件可以保持动画效果。

## 6. JPEG 格式

JPEG (Joint Photographic Expert Group, 联合照片专家组格式), 文件后缀名为 “.jpg” 或 “.jpeg”, 是最常用的图像文件格式, 由一个软件开发联合会组织制定, 是一种有损压缩格式, 能够将图像压缩在很小的储存空间, 图像中重复或不重要的资料会丢失, 因此容易造成图像数据的损伤。尤其是使用过高的压缩比例, 将使最终解压缩后恢复的图像质量明显降低, 如果追求高品质图像, 不宜采用过高压压缩比例。但是 JPEG 压缩技术十分先进, 它用有损压缩方式去除冗余的图像数据, 在获得极高压缩率的同时, 能展现丰富生动的图像。

换句话说,就是可以用最少的磁盘空间得到较好的图像品质。而且 JPEG 是一种很灵活的格式,具有调节图像质量的功能,允许用不同的压缩比例对文件进行压缩,支持多种压缩级别,压缩比率通常在 10:1 到 40:1 之间,压缩比越大,品质越低;相反,压缩比越小,品质越好。例如,可以把 1.37Mb 的 BMP 位图文件压缩至 20.3KB。当然也可以在图像质量和文件尺寸之间找到平衡点。JPEG 格式压缩的主要是高频信息,对色彩的信息保留较好,适合应用于互联网,可减少图像的传输时间,可支持 24bit 真彩色,也普遍用于需要连续色调的图像。

JPEG 格式是目前网络上最流行的图像格式,是可以把文件压缩到最小的格式,在 Photoshop 软件中以 JPEG 格式储存时,提供 11 级压缩级别,以 0-10 级表示。其中 0 级压缩比最高,图像品质最差。即使采用细节几乎无损的 10 级质量保存时,压缩比也可达 5:1。以 BMP 格式保存时得到 4.28MB 图像文件,在采用 JPG 格式保存时,其文件仅为 178KB,压缩比达到 24:1。经过多次比较,采用第 8 级压缩是存储空间与图像质量兼得的最佳比例。

## 第二节 纸质档案数字化

纸质档案数字化工作经过十几年的发展,流程已比较成熟。根据《纸质档案数字化规范》(DA/31-2017),纸质档案数字化流程包括档案出库、数字化前处理、档案扫描、图像处理、数据挂接、数字化成果验收与移交、档案入库等环节,具体见图 6-4。

### 一、档案出库

档案保管部门按照纸质档案数字化工作方案,确定的数字化对象,开展档案调取、清点、登记等前期准备工作,并提交档案出库申请,经相关责任人批准后,严格按照档案库房管理规定为数字化对象办理相关出库手续,并与数字化部门共同清点无误后,对档案进行交接出库。

在纸质档案数字化过程中,应设置距离数字化加工场所较近的保管库用于临时存放纸质档案,并对纸质档案的领取与归还进行严格管理,认真做好检查、清点、登记等工作,确保纸质档案的安全。

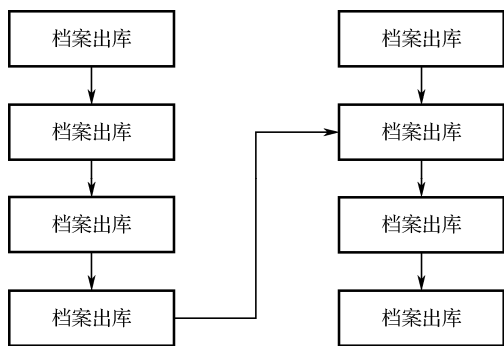


图 6-4 纸质档案数字化流程图

## 二、数字化前处理

### 1. 确定扫描页

原则上应将确定为数字化对象的纸质档案全部扫描，不宜进行挑扫。如有不需要扫描的页面应加以标注。

### 2. 编制页号

有些档案在整理时没有编页号或页号错误，应对没有页号或页号不正确的档案重新编制页号。重新编页的目的，一是方便扫描完成后档案重新装订归位，二是方便扫描时定位。重新编制页号时，应在统一位置书写页号，且不压盖档案内容。书写页号所使用的笔、墨等不应破坏档案原件或不对档案长期保存造成影响。应将破损页面、缺页等特殊情况进行登记。

### 3. 目录数据准备

按照《档案著录规则》（DA/T18）或自主确定的规则逐项输入著录项。录入时要求认真仔细，不许出现漏著、错著现象。认清节点，正确录入数据信息，发现录错节点或位置，要及时通知负责人或主管采用技术手段进行更正，无法更正的要重新补录。录入中若发现丢页、错页、漏标页号和错标页号时，应及时告知负责人，由负责人安排返工处理或退至档案所属单位重新整理。

录入完成后，要核对著录项目是否完整、著录内容是否规范、准确，发现

不合格的数据应要求进行修改或重录。

#### 4. 拆除装订

以对纸质档案的保护为原则确定是否拆除装订。如需拆除装订物，应注意保护档案不受损害，并对排列顺序不准确的档案进行重排。特殊装订，而且拆除装订后需恢复原状的档案，在拆除装订物时应采用拍照或标记等方式记录档案原貌，以便于恢复。

#### 5. 技术修复

破损严重或其他无法直接进行扫描的纸质档案，应先由专业技术人员进行技术修复。折皱不平影响扫描质量的纸质档案，应先进行压平等相应技术处理。

### 三、档案扫描

#### 1. 基本要求

档案扫描应根据纸质档案原件实际情况、数字化目的、数字化规模、计算机网络和存储条件等选择相应的扫描设备，和进行相关参数的设置和调整。参数的设置和调整应保证扫描后数字图像清晰、完整、不失真，图像效果最接近档案原貌。

#### 2. 扫描设备

扫描设备的选择应特别注意对档案实体的保护，尽量采用对档案实体破坏性小的扫描设备进行数字化。

超出所使用扫描仪扫描尺寸的档案，可采用更大幅面扫描仪进行扫描，也可以采用小幅面扫描仪分幅扫描后进行图像拼接的方式处理。分幅扫描时，相邻图像之间应留有足够的重叠，并且采用标注等方式明确说明分幅方法；若后期采用软件自动拼接的，重叠尺寸建议不小于单幅图像对应原件尺寸的 1/3。

遵循相关设备的使用规律，对扫描设备进行定期维护、保养。



### 3. 扫描色彩模式

为最大限度保留档案原件信息，便于多种方式利用，宜全部采用彩色模式进行扫描。

页面中有红头、印章或插有照片、彩色插图、多色彩文字等的档案，应采用彩色模式进行扫描。

页面为黑白两色，并且字迹清晰、不带插图的档案，也可采用黑白模式进行扫描。

页面为黑白两色，但字迹清晰度差或带有插图的档案，可采用灰度模式扫描。

### 4. 扫描分辨率

扫描分辨率的选择，应保证扫描后图像清晰、完整，并综合考虑数字图像后期利用方式等因素。扫描分辨率应不小于 200dpi。如文字字号偏小、密集、清晰度较差时，建议扫描分辨率不小于 300dpi。如有 COM 输出、仿真复制、印刷出版等其他用途时，可根据需要调整扫描分辨率。需要进行 COM 输出的档案，扫描分辨率建议不小于 300 dpi；需要进行高精度仿真复制的档案，扫描分辨率建议不小于 600 dpi；需要进行印刷出版的档案，可结合档案幅面、印刷出版幅面、印刷精度要求等选择合适的分辨率。

### 5. 存储格式

纸质档案数字图像长期保存格式为 TIFF、JPEG 或 JPEG2000 等通用格式，图像压缩率的选择可根据实际应用的需求而定。

利用纸质档案数字图像时，也可从网络浏览速度、易操作性、存储空间占用等方面进行综合考虑，将图像转换为 PDF 等其他格式。同一批档案应采用相同的存储格式。

### 6. 图像命名

应以档号为基础对数字图像命名。图像命名方式的选择应确保图像命名的唯一性。建议将数字图像存储为单页文件，并按档号与图像流水号的组合对图像命名。数字图像确需存储为多页文件时，可采用该档案的档号为图像命名。应科学建立纸质档案数字图像的存储路径，确保数据挂接的准确性。

## 四、图像处理

### 1. 图像拼接

对分幅扫描形成的多幅数字图像，应进行拼接处理，合并为一个完整的图像，以保证纸质档案数字图像的整体性。拼接时应确保拼接处平滑地融合，拼接后整幅图像无明显拼接痕迹。

### 2. 旋转及纠偏

对不符合阅读方向的数字图像应进行旋转还原。对出现偏斜的图像应进行纠偏处理，以达到视觉上基本不感觉偏斜为准。

### 3. 裁边

如需对数字图像进行裁边处理，应在距页边最外延至少2至3毫米处裁剪图像。

### 4. 去污

如需对数字图像进行去污处理，以去除在扫描过程中产生的污点、污线、黑边等影响图像质量的杂质，应遵循展现档案原貌的原则，处理过程中不得去除档案页面原有的纸张褪变斑点、水渍、污点、装订孔等痕迹。

### 5. 图像质量检查

图像不完整、无法清晰识别或图像失真度较大时，应重新扫描。对于漏扫、重扫、多扫等情况，应及时改正。数字图像的排列顺序与档案原件不一致时，应及时进行调整。对数字图像拼接、旋转及纠偏、裁边、去污等处理情况进行检查，发现不符合图像质量要求时，应重新进行图像处理。

## 五、数据挂接

可借助相关软件对数据库中的目录数据与对应的纸质档案数字图像进行挂接，以实现目录数据与数字图像的关联。

逐条对挂接结果进行检查，包括目录数据与纸质档案数字图像对应的准确性、已挂接数字图像与实际扫描数量的一致性、数字图像是否能正常打开等，发现错误及时纠正。

## 六、数字化成果验收与移交

### 1. 验收方式

档案部门应成立专门的验收组，对纸质档案数字化成果进行验收。可采用计算机自动检验与人工检验相结合的方式，对纸质档案数字化成果进行验收检验。

### 2. 验收内容

纸质档案数字化成果包括数字图像、档案目录数据、元数据、数字化工作中产生的工作文件、存储载体等。目录数据验收主要包括数据库中各条目的内容、格式等的准确程度、必填项是否填写等。元数据验收主要包括元数据元素的完整性和赋值规范性等。数字图像验收主要包括数字化参数、存储路径、命名的准确性、图像的完整性、排列顺序的准确性、图像质量等。数据挂接验收主要包括目录数据与其对应数字图像挂接的准确性等。工作文件验收主要包括工作文件的完整性、规范性等。存储载体验收主要包括载体的可用性、有无病毒等。

### 3. 验收指标

采用计算机自动检验的项目应进行 100% 检验，检验合格率应为 100%。对于无法用计算机自动检验的项目，可根据情况以件或卷为单位采用抽检的方式进行人工检验。抽检比率不得低于 5%，对于数据库条目与数字图像内容对应的准确性，抽检合格率应为 100%，其他内容的抽检合格率应不低于 95%。

### 4. 验收结论

每批纸质档案数字化成果质量检验达到规定的要求时，应予验收通过。验收未通过的，应视情况进行返工或修改，再重新进行验收。验收完成后，须经验收组成员签字。验收通过的结论，必须经相关领导审核、签字后方可生效。

## 5. 移交

验收合格的数据应按照纸质档案数字化工作方案要求及时移交，并履行交接手续，移交清单如表 6-1。

表 6-1 纸质档案数字化成果移交清单

批次					
全宗号					
内容描述					
移交数字图像数量（页）		移交条目数（条）		数字化工作文档（件、册）	
载体起止顺序号		移交载体类型、规格			
检验内容	单位名称				
	移交单位：			接收单位：	
准确性检验					
完整性检验					
可用性检验					
安全性检验					
载体外观检验					
填表人（签名）	年 月 日			年 月 日	
审核人（签名）	年 月 日			年 月 日	
单位（印章）	年 月 日			年 月 日	

## 七、档案归还入库

### 1. 档案装订

纸质档案数字化工作完成后，拆除过装订物的档案如需装订，应注意保持档案原貌，做到安全、准确、无遗漏。

### 2. 档案归还入库

按照档案入库相关要求对纸质档案进行处理和清点，并履行档案入库手续。办理档案入库手续的双方最好进行逐页清点。有的单位为了省事，不进行逐页清点，造成事后无法弥补的损失。

## 第三节 多媒体档案数字化

多媒体档案是企业档案资源的重要组成部分，根据档案数字化原则，一般应优先数字化。由于多媒体档案数字化流程及有关管理要求与纸质档案基本相同，本节主要对录音、录像档案数字化流程及技术要求作阐述。

### 一、录音档案数字化

录音档案数字化，是通过播放设备播出录音，输出模拟音频信号，模拟音频信号通过音频转换卡转换成数字信号并记录下来，完成录音的数字化。在这个过程中，最重要的是转换技术参数的设置，根据《录音录像档案数字化规范》（DA/62-2017），录音档案数字化的技术参数应满足如下要求。

- （1）采样率：不低于 44.1kMz。对于珍贵或有特别用途的档案，采样率不低于 96kMz。
- （2）量化位数：24bit。
- （3）声道：以原始声道数记录。
- （4）文件格式：WAV 格式。

录音档案数字化设备及流程如图 6-5 所示。

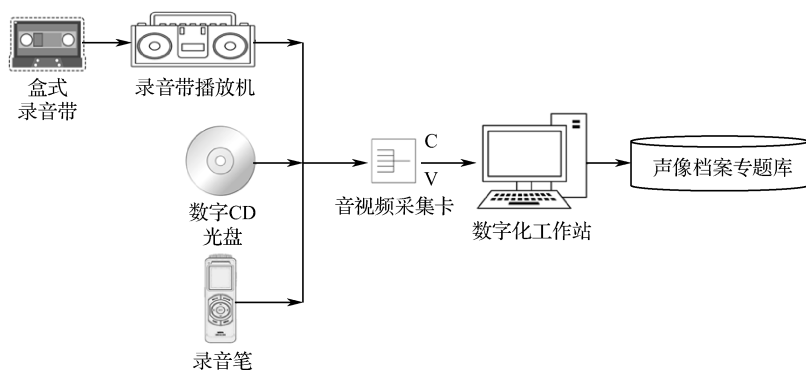


图 6-5 录音档案数字化流程图

## 二、录像档案数字化

录像档案数字化，是通过播放设备播出录像，输出模拟视频信号。模拟视频信号通过视频转换卡转换成数字信号并记录下来，完成录像的数字化。在这个过程中，最重要的是转换技术参数的设置，根据《录音录像档案数字化规范》（DA/62-2017），录像档案数字化的技术参数应满足如下要求。

（1）视频编码格式：采用 H.264 或 MPEG-2 IBP 格式。对于珍贵或有特别用途的录像档案，可采用无压缩的方式。

（2）帧率：与档案原件相同。

（3）画面宽高比：与档案原件相同。

（4）分辨率：采集为标清视频时，分辨率为 720×576（档案原件为 PAL 制式、SECAM 制式）或 720×480（档案原件为 NTSC 制式）；采集为高清视频时，分辨率不低于 1920×1080。

（5）色度采样率：采集为标清视频时，色度采样率不低于 4:2:0，采集为高清视频时，色度采样率不低于 4:2:2。

（6）视频量化位数：不低于 8bit。对于珍贵或有特别用途的录像档案，视频量化位数不低于 10bit。

（7）视频比特率：采集为标清视频时，视频比特率不低于 8Mbit/s，采集为高清视频时，视频比特率不低于 16Mbit/s。

- (8) 音频编码格式：PCM。
  - (9) 音频采样率：不低于 48kHz。
  - (10) 音频量化位数：不低于 16bit。对于珍贵或有特别用途的录像档案，音频量化位数采用 24bit。
  - (11) 声道：以原始声道数记录。
  - (12) 文件格式：AVI 或 MXF 格式。
- 录像档案数字化设备及流程如图 6-6 所示。

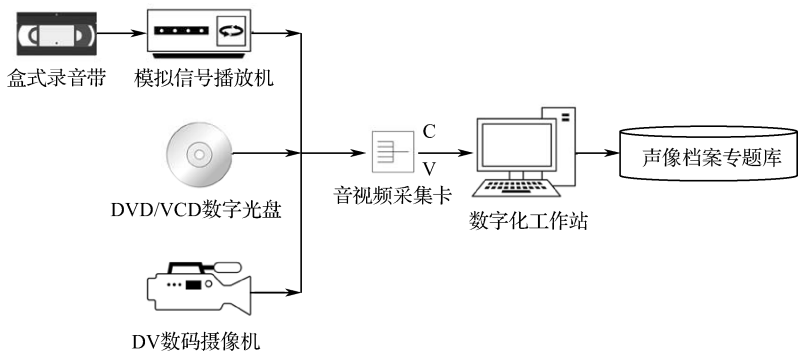


图 6-6 录像档案数字档案化流程图

## 第四节 档案数字化外包安全管理

由于存量档案数字化工作量大，在数字档案馆建设中，不少单位将档案数字化任务外包给具有一定能力的专业化公司，这是快速形成数字资源的重要方式，但在这个过程中要特别注意安全管理，采取一系列措施防止档案信息丢失、敏感和涉密信息泄露。为此，应按照以下要求做好有关工作。

### 一、外包方档案部门的安全管理工作

作为档案数字化外包工作执行部门，为保证档案数字化过程中的安全与保密，档案部门必须贯彻以下要求。

- (1) 档案部门成立由主要领导或分管领导参加的档案数字化外包管理小组，

明确档案数字化管理的部门、人员及其职责。

(2) 根据档案数字化工作规划或计划,确定数字化外包档案的范围,提出档案数字化外包安全管理要求和技术指标。

(3) 提出档案数字化外包招标文件中有关安全管理的要求,协助制定招标文件,审定合同。

(4) 对数字化服务机构的相关资质、业绩、人员、设备和加工软件等进行考察,并了解是否存在违约行为、安全事故等不良记录。在同等条件下,应优先选用具有与数字化加工相关涉密资质的数字化服务机构。

(5) 建立档案数字化安全保密制度,与数字化服务机构签订安全保密协议,并对档案数字化加工人员进行安全保密教育。

(6) 严格按照国家有关规定开展档案数字化外包服务工作,严禁将涉密国家秘密或敏感信息的档案数字化外包。

(7) 制订档案实体交接、数字化加工过程管理、数字化成果验收与交接、存储介质管理、档案实体保护等操作规程或规章制度。

(8) 建立档案数字化外包项目管理档案,记录档案部门和数字化服务机构实施档案数字化外包项目的全过程。

(9) 建立监管机制,对数字化服务机构的保密、安全措施落实情况进行监督、检查,防止档案实体受损、丢失,杜绝数字化服务机构擅自复制、留存、使用档案信息的行为。

## 二、承包数字化工作的服务机构的安全管理

为保证安全与保密,在档案数字化工作中,除档案在企业内部的安全与保密,还要对承担数字化工作的服务机构进行有效的管理,以下是企业在档案数字化工作中必须关注和落实的要点。

(1) 数字化服务机构必须具有工商管理部门核发的有效营业执照,业务范围必须包括档案数字化加工或数据处理类项目。

(2) 数字化服务机构的法人必须是中华人民共和国境内注册的企业法人或事业单位法人,股东及工作人员必须为中华人民共和国境内公民,国家另有规定的除外。

(3) 数字化服务机构的工作人员必须提供本人身份证明和公安部门提供的



无犯罪记录证明，必要时提供政审材料。

（4）数字化服务机构必须与其工作人员签订符合国家劳动法律法规要求的劳动合同。

（5）数字化服务机构的人员数量与素质、技术与管理水平、设施与设备状况能够满足拟承担项目的要求。

（6）数字化服务机构必须制订并执行数字化安全保密制度，制订并执行档案实体交接、数字化加工过程管理、数字化成果验收与交接、存储介质管理、档案实体保护等操作规范和管理制度。

（7）数字化服务机构应建立安全岗位责任制，配备专人负责安全保密工作。

（8）数字化服务机构应对工作人员进行安全保密教育和必要的上岗培训，并与工作人员签订保密协议，明确规定工作人员不得阅读、摘抄、外泄档案内容和其他安全保密责任、义务，安全保密协议应报送档案部门备案。

（9）数字化服务机构必须积极支持、配合档案行政管理部门的安全保密检查。

### 三、数字化场所的安全管理

数字化加工场所一般设在档案部门独立、可封闭的建筑内。数字化加工场所应符合防盗、防火、防尘、防水、防潮、防高温、防日光及防紫外线照射、防有害生物、防污染等安全管理要求。数字化加工场所应配备满足安全管理需要的视频监控设备，确保档案暂存处、数字化加工工位、服务器、数据导出端及门窗等无监控死角；视频监控系统应由档案部门专人负责，数字化加工场所设于档案部门之外的，档案部门应定期检查视频监控系统，数字化服务机构应将视频监控数据移交档案部门保存；视频监控数据自产生之日起保存不少于6个月；档案部门应定期对视频监控数据进行回放检查，在删除视频监控数据之前，要留存视频回放安全检查记录。数字化加工场所应配备符合国家标准并满足工作需要的档案装具，用于分别存放待数字化处理和已数字化处理的档案。数字化加工场所须切断所有档案数字化加工设备的无线网络功能，并定期进行相关检测。

数字化工作人员存放随身物品要有专用储物箱柜，并与档案装具分区放置；数字化加工场所不得有非工作需要的私人物品，包括照相机、摄像机、手机、

录音机、笔记本电脑、平板电脑等各类电子设备和各类移动存储介质；严禁擅自将数字化加工场所内的物品带离现场。工作人员要挂牌上岗，接受身份核查登记和安全检查，严禁无关人员进入数字化加工场所。工作人员不得在数字化加工场所内从事与数字化无关的活动，严禁在数字化加工区内喝水、进食、吸烟等，严禁携带火种进入数字化加工场所。档案部门和数字化服务机构应指定有关人员经常对数字化加工场所进行巡查，确保有关数字化加工和管理的各项规章制度和操作规范得到切实贯彻和执行。

#### 四、数字化加工设备、网络环境与数据载体的安全管理

在档案数字化加工过程中，建议使用档案部门提供的相关设备，使用数字化服务机构设备的，档案部门应当对其进行必要的安全检查，是否存在失泄密隐患。

档案数字化加工使用的计算机、扫描仪等设备，必须采用技术手段或专业物理设备封闭所有不必要的信息输出装置或端口，如 USB 接口、红外线、蓝牙、SCSI 接口、光驱接口等，封闭的装置或端口要定期进行检查。

推荐使用国产设备并使用正版软件。数字化安全与网络软硬件必须使用通过国家安全认证的国产品牌产品。除必要的操作系统、杀毒软件、加工软件和第三方安全管理软件外，档案数字化加工的计算机不允许安装任何与加工无关的软件。

档案数字化加工网络要与其他网络物理隔离，禁止使用无线网卡、无线键盘、无线鼠标等设备。档案数字化加工网络环境应配备具有权限管理、设备管理、端口管理、日志管理和安全审计等功能的数字化加工安全保护系统，准确记录授权用户的访问行为、设备接入和电子档案信息流向等信息。

档案数字化加工系统应具备流程定义、任务分配、过程跟踪、质量检测、成品制作、数据验收、数据备份管理等功能，并分别设置管理员、保密员、审计员，实行“三员分离”。

建议由档案部门提供计算机硬盘、移动存储介质以及确保数据能可靠清除的设备等，并逐一进行检查、登记。数字化工作完成后，这些设备必须交由档案部门统一保管或销毁，严禁擅自带走。

用于档案数字化加工的设备和存储介质，严禁与其他设备和存储介质交叉

使用，非数字化专用的设备和存储介质严禁带入数字化加工场所。

使用的移动存储介质和刻录设备应由档案部门指定专人保管，并对使用情况进行记录。档案数字化成果的拷贝和刻录应相对集中。档案部门应指定专人负责移动存储介质数量的清点，数字化服务机构完成拷贝或刻录的数据介质（包括损坏的数据介质）应及时交给档案部门指定的人员，并办理交接手续。档案数字化设备和存储介质不得擅自送外维修，必须送外维修的应办理书面审批手续，并由档案部门人员现场监督。

处理尚未开放档案的信息设备的管理和使用，应符合国家有关秘密载体管理和使用的相关规定。

## 五、档案实体的安全管理

档案部门要对拟数字化的档案进行涉密性、完整性、有序性和档案实体与目录的一致性检查。涉密档案要予以筛除，档案实体破损、残缺的要进行登记和处理，档案实体与文件目录不对应的要进行必要的记录或标示。

档案部门人员应按照工作计划分批调档，并与数字化服务机构的档案接收人员进行清点、核对，双方确认准确无误后填写档案交接清单，一式两份，注明交接档案的内容、数量、状况、交接时间和经办人等。

档案数字化加工不得损毁档案，出现档案损毁的，需按有关规定进行处罚，并进行修复和登记。需要拆装档案时，应尽可能保持档案原貌。

档案数字化过程中要建立档案流程单，流程单包括档号、加工工序、设备编号、数量、经手人、加工时间等，档案流程单应与档案实体同步流转。

发现有涉密标识且无解密标识的档案，数字化外包服务机构应停止该档案的数字化加工，在登记目录后立即将档案移交给档案部门。

正在进行数字化加工的档案必须每天入库（柜），不得在加工工位上留存过夜。数字化档案要专人专柜保管，数字化加工完毕的档案要及时归还入库。对于离库时间较长或有虫霉隐患的档案，应进行消毒杀虫处理。

## 六、档案数字化成果移交与设备的安全管理

档案数字化任务完成后，档案部门应组织专业人员按照要求，对向档案部

门移交的数字化加工介质（如存储介质、移动介质、备份介质等）、加工监控视频回放安全检查记录、档案实体出入库交接记录、加工人员变更记录等进行安全保密专项验收；凡未开展安全保密专项验收或验收不合格的，不得对项目进行总体验收。

档案数字化成果必须通过完整性、准确性、可用性和安全性检测，检测合格后双方办理数据交接手续。

档案数字化任务完成后，数字化服务机构应会同档案部门拆除其自带加工设备中的硬盘等存储介质，并将其与数字化过程中使用过的其他移动存储介质一起移交给档案部门，并办理相关移交手续。

档案数字化任务完成后，档案部门必须组织专业人员对数字化服务机构所用的设备进行检查，以确保其设备中无信息留存。凡存有信息的，必须清除信息，做安全处理。

## 第六章 电子文件归档和电子档案管理

企业数字档案馆建设中，档案资源数字化是一项重点任务。档案资源数字化中，电子文件归档是一项难度最大且十分关键的工作，它关系到数字档案馆建设的效果。通过数字档案馆建设解决本企业电子文件归档问题，日益成为当前数字档案馆建设关键。因此，本章详细阐述企业数字档案馆建设中的电子文件归档和电子档案管理，本章所述的管理功能是纳入电子档案管理系统的功能。

### 第一节 电子文件管理原则和要求

随着信息技术的不断发展，办公自动化在各级各类国家机关、社会组织、企业、事业单位中普及，计算机辅助设计、计算机辅助制造及工业互联网在科研、工业领域的成熟应用，电子商务（EC）在国内和国际贸易中的应用，使电子工作环境（Electronic Work Environment）日益普遍。很多机构的文件在计算机上产生，在网上进行处理，形成了电子文件。

电子文件的出现是一场记录革命，也引发了包括文件和档案管理在内的记录管理的革命。这场革命反映在管理技术、方法、体制、原则、理念等各个层面，不同层面之间相辅相成、互相关联、互相作用，其中任何一个方面的滞后都可能对管理目标的实现、管理系统的有效运转带来不良影响。与传统的文件、档案管理相比，电子文件、电子档案的管理具有相关因素多，各因素之间的相关度高，变量多，目标复杂，技术含量更高、更新、更快等特点，从而对管理活动提出了更高的要求，需要在科学管理的思想指导下，从政策、制度与技术不同层面共同采用新的对策。

电子文件作为文件的一种形式，具有文件的基本特征，我国文件和档案管理的一般原则，对电子文件的管理同样具有宏观的指导作用。另一方面，电子文件又有许多不同于其他类型文件的特点，在其管理活动中还应遵循一定的原

则和要求。

## 一、电子文件管理原则

### 1. 全程管理原则

根据电子文件的特点和管理要求，必须建立一个完整的管理体系，对电子文件从产生到归档，再到永久保存或销毁的整个生命周期进行全程管理。

电子文件的全程管理是一种全面的管理，涉及电子文件的流程、管理规则、管理方法以及质量要求，以建立一个涵盖电子文件全部管理活动的目标体系、程序体系和技术方法体系。

电子文件的全程管理原则，应该体现在电子文件管理体制与模式的确定、管理系统的设计和运行、管理制度的内容和执行等方面，以保证电子文件在其整个生命周期受到严密控制。以电子文件管理系统对文件流程各阶段的管理和监控为例，首先，系统应对产生于企业内部各个部门及其应用系统上的电子文件进行统一归口登记，对文件进行分类、鉴定，同时监控和记录对文件的一切操作，如文件的创建、登记、修改、审核、签署、分发，文件在半现行期和非现行期的利用等。对文件进行的任何技术操作，包括存储位置的改变、数据的转载也应该记录下来。这种管理和跟踪记录，对于证明和保证电子文件的行政有效性和法律证据性有着极其重要的作用。其次，电子文件管理系统应支持本企业的文件运转流程和业务流程，在确保文件完整、安全、准确的同时，保证文件运转和业务流程的顺畅、周密和高效。

1994年至1997年，加拿大不列颠哥伦比亚大学和美国国防部在关于电子文件真实性的联合研究项目中，提出了电子文件全程管理的框架模式，认为应由以下措施构成电子文件的管理链条。

- (1) 根据预先确定的标准格式和模板编辑文件；
- (2) 根据文件的类型和用途，使用预先确定的方法认证文件；
- (3) 根据用户权限，确定接触电子文件的权力；
- (4) 在系统中嵌入工作流程，只向有关人员呈现有关文件；
- (5) 使用磁卡、密码和指纹识别等方式，限制对某些技术的接触；
- (6) 在系统内设计审计跟踪功能，记录系统中的任何操作及其结果（如文

件被修改、删减和增添等)。

这一模式在北美地区的电子文件管理实践中得到了应用和完善。

## 2. 前端控制原则

前端控制是现代文件和档案管理理念的重要内容。前段控制原则以文件生命周期理论为基础,把文件从形成到归档,至永久保存或销毁的各个阶段看作一个完整过程。在这个过程中,文件的形成是前端,处理、归档、鉴定、整理、编目等具体管理活动是中端,永久保存或销毁是末端。

(1) 传统文件、档案管理的特征是分阶段、分环节控制,文件管理和档案管理作为两个相对独立的系统分别运行,文件、档案管理的全部目标和要求被分解到不同阶段、环节和步骤之中。

(2) 前端控制则是对整个管理过程的目标、要求和规则进行系统分析、科学整合,把需要和可能在文件形成阶段实现或部分实现的管理功能尽量在这一阶段实现。

在已经建立了电子文件管理系统的地区和机构,电子文件是在系统中生成和运转的,电子文件管理过程的前端就延伸到了系统设计阶段,前端控制的形式也部分转移到系统功能的设计之中,即尽可能把文件生命周期各个阶段的管理要求设计在系统之中,以功能合理的文件管理系统作为管理电子文件的先决条件。

前端控制是实现电子文件全程管理的重要保障,是全面、系统、优化思想的集中体现。这一原则在纸质文件和其他载体文件、档案管理中是有效的,在电子文件管理中更是至关重要。其必要性主要体现在以下两个方面。

(1) 前端控制是确保电子文件真实可靠、完整安全、长期可读的有效策略。

由于电子文件具有易流失、易更改等特性,文件内容乃至文件本身出现差错的可能性比纸质文件要大得多,其中很多差错及其造成的后果是不可挽回、不可修复、不可弥补的,即使能够恢复或部分修复,往往也需要付出很高的代价。这些差错有的来自系统的缺陷,有的来自管理的失误。

针对电子文件“稍纵即逝”的特点,防止失误、控制差错的最佳时机,应放在所有可能出现失误和差错的环节之前。把保障电子文件安全、可靠、完整、可读的措施附加于文件形成过程,可以有效防止其他管理环节对电子文件的损伤和破坏。

例如,企业将电子文件的归档范围(包括归档内容和归档版本)设计在电

子文件管理系统之中，属于该范围的电子文件一“出生”就加上归档标识，使之具有防止修改、防止删除的属性，并生成相应的描述信息，从而确保其安全留存和内容真实。

加拿大不列颠哥伦比亚大学的露西安娜·杜兰蒂教授认为，“如果形成者不从电子文件产生之时起就采取措施来保证它们的可靠性（包括内容的可靠性和文件本身作为‘文件’的可靠性）的话，未来将没有多少电子文件值得保留。”

另一位加拿大档案专家特里·库克也指出，“在电子环境中，如果档案人员不积极介入文件的形成和保管过程，文件将可能不存在，或至少可能不被鉴定、保存、编目和提供利用。”

（2）前端控制是优化管理功能、提高管理效率的科学理念。

电子文件管理系统不应是对手工管理流程的简单模拟，而应该本着更科学、更合理、更有效原则对以往的文件、档案管理进行“业务流程重构”（Business Process Reengineering），以便减少乃至消除文件、档案管理过程中各个管理环节的重复、疏漏和抵触，从而达到功能合理、效率最高。前端控制是新构建的业务流程的重要特点，为实现上述目标发挥了很好的作用。

在对电子文件生命周期全过程通盘规划的基础上，把某些分散在各个业务环节的，带有一定重复性的作业提前实施，可以有效地减少重复作业和滞后作业，最大限度地提高作业效率。

例如，凡是可能在文件形成同时生成的数据（包括描述文件内容、结构、背景、版本、文件生成环境、存在状态等各方面的信息）均在当时及时采集、存储下来，文件形成时尚无法生成的文件运转、保管、利用等方面的数据则进行实时追加。所有数据均为一次采集，终身使用。这种数据采集方式与各管理环节独立采集方式相比，工作量大为减少，有效地防止了重复采集中产生的数据误差，并且保证了用于描述文件、检索文件和系统维护所需数据的完整齐全、真实可信。又如，把电子文件的版本控制、备份措施、鉴定复审、利用权限等管理要求设计在电子文件管理系统之中，由系统自动实施，可以确保各项管理措施及时、准确、到位。

## 二、电子文件管理要求

对电子文件的管理要求，我国经历了一个调整过程。



(1) 2003 年 7 月 28 日, 国家档案局令第 6 号发布《电子公文归档管理暂行办法》提出, 电子文件归档管理的要求是“有效维护电子公文的真实性、完整性、安全性和可识别性”。

(2)《电子文件管理暂行办法》提出电子文件管理要求是“确保电子文件的真实、完整、可用和安全”, 同时规定电子文件管理的原则是统一、全程、规范标准、便于利用和安全保密的管理原则。

(3) 2002 年发布的《电子文件归档与管理规范》(GB/T18894) 提出, 电子文件归档与管理的要求是“确保其真实、完整性和有效性”, 与《电子文件管理暂行办法》的要求相比, 缺少安全性的要求。

(4) 2017 年修订的《电子文件归档与电子档案管理规范》(GB/T18894), 将电子文件的管理要求规定为“应对电子文件、电子档案实施全程和集中管理, 确保电子档案的真实性、可靠性、完整性与可用性。”

由此可见, 经过实践, 电子文件管理总的要求越来越趋于一致和明确。2017 年修订的《电子文件归档与电子档案管理规范》基本统一了我国前期实践的成果。

### 1. 真实性保障原则

电子文件的真实性是指文件内容、结构和背景信息经过传输、迁移等处理后依然保持不变, 与形成时的原始状态一致。真实性是保证电子文件行政有效性和法律证据性的基础, 是电子文件反映历史面貌, 构成社会价值, 得以作为社会记忆长久保存的前提。

电子文件真实性的维护是档案界面临的难题, 主要原因来自以下两个方面。

(1) 电子文件信息的易更改性。在电子文件形成和使用过程中, 如果不采取控制措施, 极易不留痕迹地改动文件信息。

(2) 电子文件信息的软硬件依赖性。由于软硬件平台的不断升级, 若要保证电子文件信息的可读性, 必须对文件适时迁移, 以适应新的计算机技术, 而迁移过程会造成文件某些信息的变化或丢失, 影响电子文件的真实性。

认证和维护电子文件真实性是一个极其复杂、困难的问题, 引起了国际档案界极大的关注。1999 年, 加拿大电子文件专家露西安娜·杜兰蒂组织跨越四个洲, 由若干国家和地区的研究人员组成研究小组, 开展了“电子文件真实性永久保管国际研究项目”(InterPARES)。

InterPARES 项目研究的目的是,为永久保护由电子技术生成的文件及其真实性建立一套基本的理论与方法,并在此基础上,为国际组织、国家或部门在长期保管电子文件的真实性方面制订方针政策和标准规范提供指南。

为了有效应对影响电子文件真实性永久保护问题的复杂性,达到该项目的研究目的,InterPARES 项目分为四个相关的研究领域。

- (1) 保护电子文件真实性的概念要求。
- (2) 确定电子文件作为永久保存的鉴定标准和方法。
- (3) 永久保管真实电子文件的方法、规则和步骤。
- (4) 哪些方针、策略及标准可用来长期保护电子文件的真实性,这些方针、策略及标准要考虑不同国家文化的差异和多元性。

InterPARES 项目的每个研究领域都由专门的任务小组组成,小组成员来自加拿大、美国、英国、爱尔兰、瑞典、荷兰、法国、意大利、葡萄牙、澳大利亚、中国及其香港的政府机关、私人组织和企业,涉及档案科学、古文书学、凭证及版权法、计算机科学、计算机工程技术、历史、存储介质的化学性质、法律制裁等多个学科领域。

InterPARES 项目提出的解决方案,包括事前控制、跟踪记录和事后审查等阶段性措施;针对现行文件、半现行文件和非现行文件的对象性措施;基于政策、方针、程序、制度、标准的管理框架和基于系统软硬件环境、加密、签署、认证的技术框架等,从不同角度就电子文件的真实性认证和维护提出了对策。

InterPARES 项目提出了维护电子文件真实性的四项措施,如下:

(1) 确认不同类型的电子文件中,哪些要素对于保证电子文件的长期真实性至关重要,以便对这些要素加以记录和保护。为此,研究人员以文件解构定义法建立了一个“文件分析模板”,确认并定义了电子文件的全部组成要素,在不同国家通过案例研究验证和测试这一模板,并从中选择和确认能够实时和长期证明电子文件真实性的要素。

(2) 将电子文件中用户不可视的部分变成可视的,并通过与文件的智能链接,使可视性得到固定。这里不可视的部分主要指元数据,或是决定文件运行质量的物理元素等。

(3) 在上述措施不能奏效时,考虑是否有可能将该文件转换成非数字式文件。

(4) 对文件的迁移采取自我认证和自我记录的办法进行不间断的物理管理。研究人员认为,当文件仍被日常工作中所需要时,形成者可以通过建立电子文

件形成和保管方面的控制程序来确认迁移文件的真实性，当长久保存的电子文件不再被其形成者的日常活动所需要，迁移过程由一个与文件的内容或存在无关的机构来执行时，迁移的结果应由中立机构进行检验和认证。这种中立方可能是档案机构、公证人或任何具有认证职能的机构。该机构依据对整个迁移过程的完备记录，对迁移文件的真实性加以确认并做出声明。

在认定和维护电子文件的真实性问题上，国际档案界的有关研究成果为我国提供了有价值的借鉴。我国文件、档案研究和管理人员还应该根据本国的社会、政治、文化、法律背景，以及文件、档案管理体制和传统，研究并提出我国的对策和方案。

## 2. 可靠性保障原则

电子文件的可靠性概念，在国际和国内稍有差别。国际上认为，电子文件的可靠性是“可作为所要印证的事项处理完整和准确的表现形式，可以信赖，并可在后续的事项处理过程中作为依据”。

我国《电子档案管理基本术语》(DA/T58-2014)将可靠性定义为“电子文件、电子档案的内容完全和正确地表达其所反映的事务、活动或事实的性质。”

尽管国内和国外对可靠性定义不同，两者在本质上是一致的，即指电子文件的正确性，即非错误的。

电子文件的可靠性与电子文件的真实性是有区别的。真实性是指其与本来面目一致，强调的是原始性，非篡改性。而可靠性是指其表达的正确性，与是否原始是有区别的。相应的电子文件是原始状态，但并不一定能表示其本来应该表达的意思。

例如，一份没有被更改过的电子文件使用非原生软件打开后，有些内容无法表示出来，或是表示内容不完整，那么该电子文件是不可靠的。可靠的电子文件肯定具备真实性，而真实的电子文件不一定可靠。

## 3. 完整性保障原则

电子文件的完整性包括下两个方面的含义。

(1) 作为记录社会活动真实面貌的、具有有机联系的电子文件，以及其他形式的相关文件数量齐全。

(2) 每一份电子文件的内容、结构和背景信息没有缺损。

完整性是电子文件价值的重要保障,残缺不全的文件留给后人的是残缺不全的社会记忆。因此,维护电子文件的完整性是文件、档案工作人员的重要职责。

为了确保相关电子文件的齐全完整,必须掌握电子文件的形成规律和分布状况,列出相关电子文件的清单,通过系统功能和人工监控将具有有机联系的电子文件收集齐全,特别要注意在不同应用系统中分散形成的、不同媒体的、通过非正式渠道传递的、具有内容相关性的电子文件的收集和捕获。

例如,有关同一事件的电子文件一部分在电子邮箱中,一部分在电子表格或相关数据库中,将这些电子文件收集齐全,并保持有机联系才符合完整性要求。此外,还应考虑信息技术的应用并未涉及的机构业务活动领域。在同一业务中,除了电子文件外,有时还会生成其他形式的重要文件,如纸质文件、缩微胶片等。为保持这些文件之间的历史联系,确保同一活动中形成的记录信息完整无缺,需要将其一同归档。有条件的单位可将这些文件做数字化处理,作为电子文件归档和保存。

加拿大学者露西安娜·杜兰蒂认为,只有保存了这些文件之间的联系,才能保证一百年以后的读者能够看懂与同一事件相关的所有文件,了解事情的来龙去脉。因此,为了确保每一份电子文件的完整性,应分析电子文件的构成要素,并根据分析结果,形成电子文件的元数据模型。

InterPARES 项目的研究成果认为,电子文件的基本构成要素包括以下几项:

① 载体 (medium),指承载信息并具有实体形态的介质。

② 内容 (content),指文件记载的信息。

③ 物理和智能格式 (physical form&intellectual form),指合乎通信要求的文件表示规则。物理格式指决定文件外观的一组属性,包括字体、插入符、颜色、语种、数字签名、时间戳等;智能格式指表示文件生成活动、文件编制及管理背景的所有属性,文件内容的表示类型,如文本、图形、图像或混合形式,以及文件各部分内容及其排列顺序、注释等)。

④ 活动 (action),指签发文件的活动,即以文件为手段,开始、维护、改变和结束某项工作的行为。

⑤ 人员 (persons),指利用文件开展工作的实体。

⑥ 档案链 (archival bond),指同一业务活动所形成的所有文件之间的联系。

⑦ 背景 (context),指文件产生的司法、行政、业务活动以及技术环境。

在传统环境下,文件的各个要素是可视的,且互相融合、彼此渗透。在计

算机系统中, 这些要素完全可能独立存在, 分别管理, 而且有些要素是不可视的, 如对文件格式的控制命令等。

电子文件的完整性是上述电子文件基本构成要素的有机结合。为保证和维护文件的完整性, 需要有意地将这些要素集中起来。电子文件具有软硬件依赖性, 因而归档文件的支持软件及软件的文档、表达电子文件内容的基本格式及有关元数据都应属于归档范围。

#### 4. 可读性保障原则

电子文件的可读性是指文件经过存储、传输、压缩、加密、媒体转换、迁移等处理后能够以可以识读、理解的方式输出, 并保持其内容的真实性。电子文件的可读性是其存在和保存价值的基础, 如果文件不能顺利读出, 文件中的信息便成为“死信息”, 再有价值的也失去了存在的意义。

电子文件读不出来的例子, 在人们日常的工作和交往中常常遇到, 已经归档的重要的电子文件也有读不出来的情况。例如, 德国联邦档案馆中保存了近千万份民主德国政府的电子文件和数据库, 包括行政文件、农业文件、劳工统计文件、监狱登记册和党的干部人事文件等。由于产生这些文件的计算机系统资料已经丢失, 软件编码也不清楚, 存储介质已经过时, 使得这些文件无法读出。专家认为, 这意味着这些民主德国政府的文件已经丢失。

加拿大电子文件专家露西安娜·杜兰蒂教授认为, “过去十年产生的信息远远超过了人类历史上任何一个时代。但具有讽刺意味的是, 信息时代产生的绝大部分信息比任何时代都难以提供利用。”

造成电子文件读不出来的原因很多。例如, 电子文件的系统依赖性使得每一份文件只有在相应的系统环境中才能读出, 系统不兼容往往导致文件无法读出。又如, 与电子文件产生、管理和存储有关的计算机软件 and 硬件更新换代, 新的系统环境因不接受过时的文件格式而无法读出原系统生成的文件。此外, 病毒感染、载体损伤、信号衰减、加密文件无相应密码等原因也都会影响文件的顺利读出。

保证电子文件可读性的措施应该贯穿于全部电子文件管理工作的始终。

(1) 归档时应根据电子文件的类型和特点注明文件存储格式、软硬件环境、相关的数据及参数等。

(2) 通用软件产生的电子文件, 应同时收集其软件型号、名称、版本号和

相关参数手册、说明资料等。

(3) 专用软件产生的电子文件，原则上应转换成通用型电子文件，不能转换时应将该软件与文件同时归档。

(4) 对于加密电子文件，应解密后再归档，确实需要以加密方式保存的，应将其密码同时归档。

(5) 对电子文件的鉴定，要从内容和技术状态两个方面进行分析，可读性检测是技术分析的内容之一。《电子文件归档与管理规范》(2001年修改稿)中规定，归档的每套载体均应接受检验，合格率应达到100%。

(6) 要定期对脱机保管的电子文件进行抽样读取检验，发现问题要及时采取恢复措施，要根据软硬件升级换代情况适时对电子文件进行迁移作业等。

### 三、电子文件管理环节

根据电子文件的特点和管理要求，必须建立一个完整的管理体系，对电子文件从产生到永久保存或销毁的整个生命周期进行全程管理。

电子文件管理环节是电子文件管理体系的重要内容。就电子文件管理流程的设计，国内存在以下两种观点。

#### 1. 建立在“大文件”概念基础之上的电子文件管理环节模型

建立在“大文件”概念之上的电子管理环节由中国人民大学信息资源管理学院电子文件研究中心专家和学者提出。

(1) 建立在“大文件”概念之上的电子管理环节，注重电子文件生命周期内各个阶段所有管理活动和管理要素的统筹兼顾，强调各项管理内容和要求的无缝衔接、系统整合和总体效应。

(2) 建立在“大文件”概念之上的电子管理环节，注重充分利用管理系统的各种资源，以知识管理的理念，追求系统软硬件资源和信息资源的最大共享和最大效益。

因此，建立在“大文件”概念之上的电子管理环节，对电子文件生成、流转、利用、保管等每一项具体管理内容的实施过程进行监控，便于及时发现和纠正失误，不断调整管理策略。在过程管理中，所有有助于说明电子文件重要属性和有效管理过程的信息都被记录在案，以证实电子文件在管理系统中的运

转状况，确保电子文件的管理质量。

2. 简化的电子文件管理环节模型

部分档案部门专家和学者基于档案管理的工作模型对电子文件提出了相对简化的电子文件管理环节模型。

简化的电子文件管理环节模型，根据我国档案工作理论，将电子文件的创立、登记、修改、审核、签署、分发等过程统一为形成办理环节，将电子文件电子管理过程设计为形成、捕获、归档、电子档案保管和处置等环节，其流程模型如图 6-1 所示。



图 6-1 简化的电子文件管理环节流程模型

在简化的电子文件管理环节流程模型中，有些环节又包括若干子环节。例如，归档环节还包括收集、鉴定、整理、赋权等环节，如图 6-2 所示。

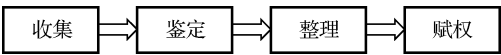


图 6-2 归档环节的子环节流程模型

第二节 电子文件形成的管理

电子文件管理的目的，一是提供业务办理，二是归档保存以备查利用。因此，电子文件形成管理至关重要，必须满足业务办理的需求和后期档案管理的需求。由于不同的业务对电子文件管理的需求不相同，本章主要是从档案的视角提出电子文件形成的要求。从档案的视角，要求电子文件需要满足真实、可靠、完整、可读四性及可归档，为此，电子文件形成应符合一定的要求。电子文件的管理要求，一般是通过数字档案馆的电子档案管理系统功能进行落实的。

## 一、保证电子文件真实、可靠、安全

电子文件形成时，应采用技术和管理相结合的方式，保证电子文件形成过程中操作的不可抵赖性、数据存储的完整性、用户身份的真实性，并妥善管理在运行过程中使用的密码、密钥等认证数据，保证电子文件不被非法更改。

(1) 制定相应的风险控制策略，防止形成电子文件的业务系统在运行和业务处理过程中发生有意或无意的危害，导致信息完整性和真实性发生变化，并具备有效的业务容量、业务连续性计划和应急计划。

(2) 电子文件形成系统，必须确保形成后的电子文件保存完整且不被修改，对电子文件进行的任何非常规的改动均能记录在相关的元数据中；能维持电子文件及其元数据的技术、结构和关系的完整性。

(3) 对电子文件的访问能够进行控制，即能够根据用户角色，限定某些电子文件使用的权限，并严格执行系统的监管控制。

(4) 电子文件安全能够得到有效控制，电子文件安全等级一旦设定，未经授权不会被改动和超权限访问。

(5) 电子文件的任何移动能够被追踪。

## 二、按一定格式存储

电子文件形成时，其存储的格式应符合长期保存要求，或可转换成长期保存格式。为保证电子文件的顺利归档，电子文件存储应尽可能采用符合归档要求的数据结构和文件存储格式作为系统运行时的存储格式，如确无法采用符合归档要求的格式，应在系统实施时所运用的数据结构和文件存储格式进行归档格式可转换性评估。电子文件无法转换成符合电子档案存储要求格式的业务系统应谨慎使用。

符合长期保存要求的电子文件存储格式选择和确定一般如下。

(1) 数据库记录电子文件、文字处理系统形成的电子文件，一般选用版式为归档存储格式；版式存储格式按照 DA/T47 的原则确定；符合 DA/T47 要求的版式电子文件存储格式标准较多，应首先选择具有国内自主知识产权的标准，在确定无国内自主知识产权标准的情况下，可选择国外开放的格式标准。有些



企业对电子档案具有统计分析要求,可采用版式、XML 两种格式同时存储的方式进行,也可将具有统计需求的字段值存储为扩展元数据。当前采用较多的版式文件存储格式有 PDF 和 OFD,此种格式当前国内使用较多。

(2) 图像电子文件的归档存储格式可根据 DA/T47 的原则和 GB/T 18894 的规定,并参照 DA/T32 的要求确定。

(3) 图形类电子文件的归档存储格式可根据 DA/T47 的原则和 GB/T 18894 的规定,并参照 DA/T32 的要求确定;其中,CAD(计算机辅助设计)系统形成的电子文件其源存储格式标准一般不公开,此类电子文件归档时,可采用源格式和交换文件格式同时存储,以尽量减小长期可读风险,交换文件格式可供选择有 STEP、DXF 格式。

(4) 由于技术的进步,国家将不定期发布有关标准和规范,各企业应及时关注,尽可能采用国家规范的格式。

(5) 电子文件在形成时,如无法采用档案管理要求的格式,应转换成符合档案管理要求的格式。确实无法转换的,应将相应阅读、显示软件系统及其软件运行需要的专用硬件与电子文件同时保存。

### 三、捕获相应的元数据

元数据是电子文件的一部分。元数据元素有些需要在电子文件形成时形成并捕获。为形成完整的元数据,对于内部信息系统中形成的电子文件,系统开发与实施时应对应元数据捕获节点与内容进行规划,将应由业务系统捕获的电子文件元数据全部形成并捕获。对于外部接收的电子文件,应将电子文件本身和元数据一并接收。

根据《文书类电子文件元数据方案》(DA/T46)定义的元数据项,元数据来源主要有业务系统(电子文件的形成系统)和电子档案管理系统。本着最便利原则和前端控制原则,文书类电子文件元数据捕获节点可做如下规划。

(1) 下列元数据应在业务系统(电子文件生成系统)捕获。

题名、并列题名、副题名、说明题名文字、文件编号、责任者、日期、文种、紧急程度、主送、抄送、密级、保密期限、件数、页数、语种;格式信息、计算机文件名、计算机文件大小、文档创建程序、信息系统描述、数字化对象形态、扫描分辨率、扫描色彩模式、图像压缩方案;电子签名数据(包括签名

规则、签名时间、签名人、签名结果、证书、证书引证、签名算法标识)。

(2) 下列元数据应在电子档案管理系统中生成或捕获。

档案馆名称、档案馆代码、全宗名称、立档单位名称、全宗号、目录号、年度、保管期限、室编案卷号、馆编案卷号、室编件号、馆编件号、文档序号、页号、摘要、当前位置、脱机载体编号、脱机载体存址、缩微号。

(3) 下列元数据在业务系统和电子档案管理系统任何一方形成或捕获均可。

主题词、关键词、人名、机构或问题、类别号、分类号, 知识产权说明、授权对象、授权行为、控制标识。

(4) 下列元数据可能会在业务系统和电子档案管理系统均有形成并捕获。

机构和人员元数据(包括机构人员类型、机构人员名称、组织机构代码、个人职位), 业务实体元数据(包括业务状态、业务行为、行为时间、行为依据、行为描述), 实体关系元数据(包括实体标识符、关系类型、关系、关系描述)。

其他类型电子文件和电子档案元数据的形成捕获节点可参照上述方法规划。

元数据也有必备项和选择项, 对于标准中规定的必备项, 在管理过程中必须形成和捕获。

## 第三节 电子文件归档

### 一、电子文件归档范围

#### 1. 电子文件归档范围界定依据

电子文件在形成并办理完毕后就应该归档。但根据档案管理规定, 一个机构和企业并不是所有的电子文件都需要归档, 而是要根据保存价值来判断是否需要归档, 只有那些具有保存价值的文件才需要归档。因此, 首先要划定电子文件的归档范围。

归档范围就是以条目或表格的形式列出需要归档的文件名称。在档案工作中, 归档范围一般与保管期限表合二为一, 在列出哪些具体文件需要归档的同时, 也列出其保管期限, 称为保管期限表。而仅归档范围来说, 一般用条目的方式列出文件的名称或类名。

国内尚无专门规范电子文件归档范围的规章制度和标准规范，因此，当前电子文件的归档，主要是参考纸质等传统载体档案。可作为界定电子文件归档范围的参考依据主要有以下几点。

#### （1）法律及规章

##### ① 《企业文件材料归档范围和档案保管期限规定》

《企业文件材料归档范围和档案保管期限规定》（国家档案局令第 10 号），对企业文件材料归档范围进行了界定，特别是对管理类（文书类）文件材料的归档范围进行了详细的界定。

##### ② 《会计档案管理办法》

《会计档案管理办法》（财政部、国家档案局令第 79 号），对会计核算文件材料归档范围作了界定。其中第六条界定了会计文件材料的归档范围为：会计凭证，包括原始凭证、记账凭证；会计账簿，包括总账、明细账、日记账、固定资产卡片及其他辅助性账簿；财务会计报告，包括月度、季度、半年度、年度财务会计报告；其他会计资料，包括银行存款余额调节表、银行对账单、纳税申报表、会计档案移交清册、会计档案保管清册、会计档案销毁清册、会计档案鉴定意见书及其他具有保存价值的会计资料。

##### ③ 各种行政法规

除国务院部门规章对文件材料的归档范围进行了界定外，还有一些行政法规也对文件材料归档范围提出了要求，这在电子文件归档中要特别注意。

2003 年 12 月 27 日第十届全国人民代表大会常务委员会第六次会议通过修订的《中华人民共和国商业银行法》的下列条款均提出了档案管理要求。

第五十一条，商业银行应当按照国家有关规定保存财务会计报表、业务合同以及其他资料。

第六十二条，国务院银行业监督管理机构有权依照本法第三章、第四章、第五章的规定，随时对商业银行的存款、贷款、结算、呆账等情况进行检查监督。检查监督时，检查监督人员应当出示合法的证件。商业银行应当按照国务院银行业监督管理机构的要求，提供财务会计资料、业务合同和有关经营管理方面的其他信息。

类似的法律还有 2009 年 7 月 8 日国务院第 73 次常务会议通过的《中华人民共和国食品安全法实施条例》中的有关内容，如下：

第二十二条，食品生产经营企业应当依照食品安全法第三十二条的规定组织职工参加食品安全知识培训，学习食品安全法律、法规、规章、标准和其他

食品安全知识，并建立培训档案。

第二十四条，食品生产经营企业应当依照食品安全法第三十六条第二款、第三十七条第一款、第三十九条第二款的规定建立进货查验记录制度、食品出厂检验记录制度，如实记录法律规定记录的事项，或者保留载有相关信息的进货或者销售票据。记录、票据的保存期限不得少于2年。

第二十八条，食品生产企业除依照食品安全法第三十六条、第三十七条规定进行进货查验记录和食品出厂检验记录外，还应当如实记录食品生产过程的安全生产管理情况。记录的保存期限不得少于2年。

第二十九条，从事食品批发业务的经营企业销售食品，应当如实记录批发食品的名称、规格、数量、生产批号、保质期、购货者名称及联系方式、销售日期等内容，或者保留载有相关信息的销售票据。记录、票据的保存期限不得少于2年。

第三十条，国家鼓励食品生产经营者采用先进技术手段，记录食品安全法和本条例要求记录的事项。

《中华人民共和国食品安全法实施条例》中的上述条款不仅对食品生产、经营过程记录（即档案）范围提出要求，而且还对保管期限提出了要求。

机构在进行电子文件归档时，均应参照并遵守这些规定，将这些电子文件纳入归档范围并归档。

## （2）规范性文件

国务院部门规章界定的文件材料归档范围只是部分，大量文件材料的归档范围是通过规范性文件来界定的，这些规定主要见诸于各类档案管理办法。

《科学技术研究档案管理暂行规定》（国档发[1987]6号，1987年国家科委、国家档案局联合下发）对科研开发类文件材料归档范围进行了界定。

## （3）标准规范

《国家重大建设项目文件归档要求与档案整理规范》（DA/T 28）对建设项目类文件材料归档范围进行了界定。

《设备仪器档案管理办法》中规定了设备仪器类文件材料归档范围。

## 2. 电子文件归档范围界定方法

长期以来，国内对电子文件管理尚缺乏深入的研究，对电子文件的归档范围大都用“参照纸质”的字样来规定。然而，机构业务信息化之后，所产生的

电子文件在内容上与非信息化时期有时存在较大变化，如果此时仍将归档范围用“参照纸质”来界定，将使很多新产生的电子文件不能纳入归档范围而无法归档。

并非一个机构形成的所有电子文件都有必要作为档案予以记录。因此，机构必须对信息化状态下所产生的电子文件进行系统的梳理，然后根据其价值进行归档范围的系统梳理，这一过程由图 6-3 简要勾画。

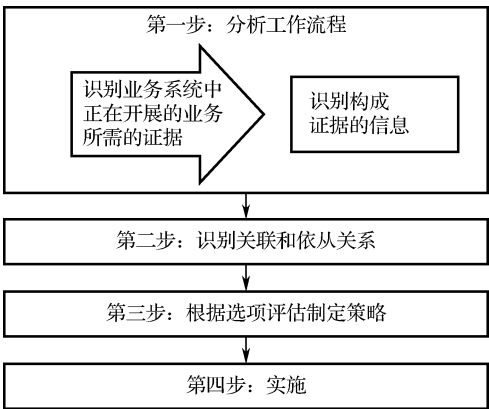


图 6-3 确定文件管理要求的步骤

（1）分析工作流程

典型的业务系统储存了大量频繁更新的数据。正因为如此，可能很难判定系统中究竟哪些信息需要作为文件予以管理，以期为业务流程提供数据。

业务系统中可包括如下：

- ① 与系统相关联并受其控制的数据元素集合（或结构化数据），例如一个数据库的条目。
- ② 由系统控制、具有明确定义数据格式的单独数字对象（或未结构化/半结构化的信息），例如文档、电子邮件或者电子表格。
- ③ 以上两者的结合。

识别文件流程伊始，必须跳出 IT 系统本身，先分析工作流程，包括相关的规章和业务需求，以确定哪些证据需要保存。

由于文件与业务流程直接关联，识别文件要借助于标准的业务流程分析技术和工具，比如，活动图表、流程分解和流程图。

在这一过程中，与机构的文件管理专业人员密切合作非常重要，因为这项

工作大部分在制定计划处置期限表之际可能就已做过。

识别文件的过程涉及以下两项主要工作：

① 识别业务系统中正在开展业务所必需的证据；

② 识别记录这些证据的信息，即“文件”。

(2) 识别作为业务证据应满足的要求

① 确定在业务系统中（全部或部分）履行的主要业务职能和开展的具体业务活动及事项处理。

这里可能要顾及的方面包括：业务流程文档说明、系统输入与输出、相关政策与程序。在高度集成的环境中，还须对多重系统进行分析，以便获得对业务流程或业务活动的完整认识。特别是在政府环境中，系统可能由多个机构共享。

② 对于由系统实施的每一项职能、活动和事项处理或业务流程，要考虑哪些证据需要由机构保存。

这些要求可能来自多个方面。要考虑诸如以下的问题：

(A) 是否有法律义务记录某些证据，一些法规可能隐含或明确规定了需要生成某些特定类型的文件。

(B) 是否有必要遵守并需要证明其合规性的正式监督文件，比如强制性标准、业务法规等。

(C) 机构是否有要求对证据予以记录的规章，例如政策、业务守则、报告制度等。

(D) 为支持针对业务流程自身所做出的决策，或将来为决策者提供信息支持时都需要哪些证据。

(E) 机构的哪些业务流程自身所做出的决策，或将来为决策者提供信息支持时需要哪些证据。

(F) 机构的哪些业务职能或活动被视为高风险或易于引起大量诉讼，因而要求更高等级的证据记录。

(G) 各种不同利益相关方是谁，他们对保存什么样的证据有哪些期望。

(H) 业内对工作流程的证据有哪些期望。

这一过程可能需要广泛咨询高级管理人员，并得到他们的确认。在实现这些目标时，有用的参考资源包括：国际标准/技术报告《文件工作流程分析》(ISO/TR 26122-2008)和澳大利亚《设计实施文件管理系统(DIRKS)手册》等。

③ 识别记录证据的内容及其相关管理信息。

并非所有包含在业务系统中的信息均有必要作为证据予以记录。因此，针对每一个所要求的证据，需要识别构成该数据的内容或数据。

在管理单独的数字对象系统中，以字处理文档为例，其数据已经汇集为一个具有逻辑结构的文档。也就是说，识别具体的、含有可作为某一特定业务活动或事项处理证据内容的文档或报告可能相对容易。

对于其他系统，就需要分析系统中的数据结构、数据模型和分类模型，识别出共同构成证据内容的具体数据元素，从而提供所必需的证据。

重要的是，要注意构成证据的内容或数据或许并不仅仅存在于自身系统中，它也可能存在于其他系统中，存在于有关系统、程序和纸质文件等文档中。特别是在高度集成的环境中，部分所需证据可跨越多个系统保存，而一些系统或构件可能与其他机构共享。

可能存在大量不同的、可用于构成证据内容的元素，应在业务需求和风险评估的基础上，决定哪些内容最适合构成所需证据。文件要充足，即对开展的业务活动或事项处理应该有充分的证据，以便能够对该行为做出说明。因此，对于主要的初始性活动要做详细记录，而对于常规低风险的操作，则可只记录最低限度的识别信息。

图 6-4 提供了一个由业务系统控制的数据库内容示例。在此示例中，文件由一组来自许多不同字段的相关数据元素构成，每份文件由数据库中业经识别的数据元素以及链接这些元素并提供支撑该文件的必要结构和背景所需的相关元数据组成。

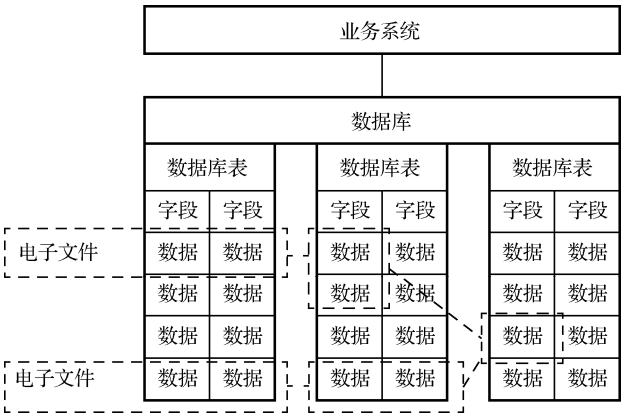


图 6-4 数据库中某一电子文件的信息构件/数据元素示例 1

图 6-5 描述了一个最简单的数据库表示例，是某一人力资源管理系统关系数据库的一部分。每个数据库表是数据库的一部分。该数据库包含相互密切联系的信息。数据库 A、B、C 分别提供了有关人员、工资和成本中心的数据。

每一个数据库表由许多列构成，代表包含数据元素的字段。每个数据库表中的行建立了不同字段中数据元素之间的联系。图 6-5 中有许多潜在的电子文件。这些文件由大量相互关联的数据元素予以表达，这些数据元素可跨越一个或多个数据库链接在一起，并由来自一个或多个字段的数据元素构成。

数据库表 A：人员			
员工编号	姓名	地址	城市
SA001	李晓明	交通大学路 1 号	北京
SA002	张晓明	交通大学路 2 号	北京
SA003	王晓明	交通大学路 3 号	北京
SC004	黄晓明	清华东路 2 号	北京
SD005	郑晓明	交大东路 51 号	北京

数据库表 B：工资			
支付代码	级别	年限	工资额
A6547	A	1	5600
A6548	A	2	6500
A6549	A	3	8000
C7475	C	8	12000
B4563	B	5	10000
C7478	C	9	13000
B6968	B	6	11000

数据库表 C：成本中心	
员工编号	支付代码
SA001	A6547
SA002	A6548
SA003	A6549
SC004	C7475
SD005	B4563

图 6-5 数据库中某一电子文件的信息构件/数据元素示例 2

从该系统中可识别出员工人事信息、工资和成本数据元素。可以认为，包含在数据库表 B 中的信息并未构成一份文件，而仅仅是员工工资文件的一部分。这是由于数据库表 B 所包含的数据是补充性的，只有置于数据库表 A 的一个职员背景之中，才可获得作为一份文件构件的值。数据库表 B 的信息本身很可能来自外部文件，如某一工作单位的合同。

还应注意的是，在某些情况下，数据库中所确认的文件之间可能有重叠。



在关系数据库中，构成一份文件的部分数据元素，也可能构成由同一数据库生成其他文件的一部分。一旦构成电子文件的数据元素之间发生重叠，业务系统必须有能力确保相关的两份电子文件在各自的保管期满之前，不得将共享的数据元素销毁。

#### ④ 识别关联及依存关系。

文件的一个关键特性，就是不能孤立地去理解。为了提供文件的背景，可能需要有关工作流程或业务系统的附加信息。因此，除了识别文件关键数据要素，还需要识别所需附加信息等相关管理信息，以便长久管理证据内容。

为了保证文件能够得以理解，验证其作为证据的可靠性，或考虑将来文件需要从一个系统移至另一个系统。需要的系统信息可包括位置、系统问题或故障、规模、实施的业务规则、文件格式、安全、隐私管理、数据结构、数据和分类模型、工作流的路由规则和审计追踪等。

有关工作流程所需的信息，可能包括相关的政策和程序文档，以证明做出的决策和运行的流程符合已审定的标准。此外，在实际中很多业务流程会延伸到单一业务系统之外。因此，在制定业务系统的文件管理策略之前，与其他系统或以纸质形式存在的相关资料的必要联接，也必须予以考虑。

相互依存的关键，在于文件需要保存多久。文件必须保存一段时间，以符合授权立法机关和司法管辖区域的要求以及业务需求。有关文件必须保存多久的决定，应在处置期限表中予以说明。机构需要满足相关司法管辖区域管理部门文件保管和处置的要求。

需要保存更长时间的文件，通常需要更为严格的控制，以确保这些文件得到有效管理，并在要求的期限内可以持续访问，同时，在已批准的处置期限表中予以规定。依据对此前文件访问的需要而定，机构可决定不把所有文件保存在现行系统中。最重要的，是要符合业经认可的维护标准，以便能够识别和检索。

文件元数据是文件不可分割的有机组成，可用于控制文件维护时间的长短、设定访问权限和访问限制，便于文件搜索与检索。生成、捕获和管理文件元数据为文件的识别、理解和检索提供了可能，对保护其证据的真实性、可靠性和完整性来说至关重要。元数据应在符合已认定的文件元数据标准的前提下予以捕获，同时也要符合司法管辖区域和机构规定的要求。

元数据并不需要与文件内容一起保存，只要它们以某种形式得以联接或关联即可。元数据可考虑保存在业务系统以外的系统中，或在文档或工具（如 XML 模型及数据）之中，或使文件易于理解并能长期有效保存的分类模型之中。

特别是在数据库环境中，可能很难区分文件的内容及其元数据。例如，提供某个特定人员在某一特定日期或时间，访问某份文件证据的元数据，它本身就是一份文件。在业务系统内的元数据常常作为一个整体从属于该系统，即从整体上应用于系统中所有的文件，而非单份文件。它可驻留在系统规则文档或系统文档中，而非应用于单份文件。

关于文件的“归档”、保管和处置，在文件管理和信息技术领域中，“归档”有着不同的含义。将数据“归档”至第二层级或脱机储存都不会改变文件管理的要求，而且也不认为这样做就满足了文件保管和处置的要求。此外，为业务的连续性或灾难恢复目的所进行的业务系统备份，也并未满足处置要求。

⑤ 基于选项评估，制定核心文件管理流程的策略。

遵循识别文件形式的证据及其相关的依附关系和关联要求，制定管理文件的恰当策略。这些策略必须基于与文件相关的风险评估。

考虑到证据的真实性和可靠性，内容必须固化到某一时间点上，并且不可更改。由于业务系统通常包含动态的、接受定期更新的现行数据，因此，必须执行保存“固化”文件的策略。这些策略会受到由哪个系统保存文件之决定的影响，可通过选项评估掌握相关资料。

在采用功能要求之前，机构需要考虑借助业务系统应用程序自身的内部机制，能够提供何种程度的文件管理功能，或者是否可通过与系统外部的、有能力提供必要文件管理功能的应用软件的交互作用，来满足要求。

某单位的强制性功能要求列出了必须加以关注的核心文件管理流程。如图 6-6 所示，实施这些要求的选项可包括以下方面。

- 设计执行内部文件管理功能的业务系统；
- 与已经确认使用的文件管理系统（例如电子文件管理系统）进行整合；
- 在业务系统中建立导出功能，直接将文件及其相关的元数据导出到某一已经确认的文件管理系统。

图 6-6 中所描述的选项并不全面，机构可根据实际情况，选择其他适合的业务系统文件管理方案。

对于那些管理单独数字对象的业务系统来说，“固化”一个文件，可以通过系统控制得以实现，如设定对象为“只读”，并始终运用文件元数据来记录文件的管理和使用。

相反，通常数据库系统包含的数据是频繁更新的、可操作的、非重复或实时的，因而对确保文件的“固化”造成困难，解决这一挑战可有以下策略。

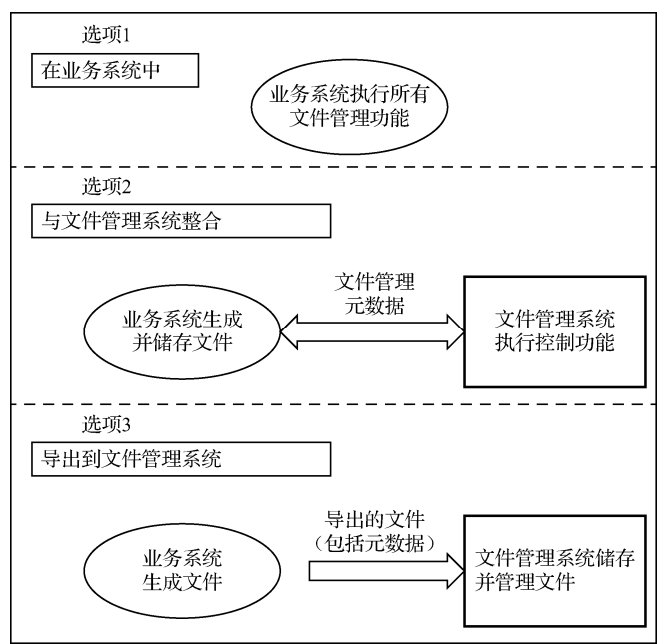


图 6-6 管理业务系统中生成文件可采用系统选择方案选项

策略一：设计防止将存入系统的特定数据覆盖或删除的控制机制。例如，准许数据的更新，但要在历史状态的字段里记录更新之前的数值。文件是特定字段与相关事件历史信息结合而藏，但这并不意味着系统内所有的数据更改都需要保留，只有那些被认定构成证据内容所必需的数据元素才需要保存。

例 6-1 一名员工将一个新客户的详细信息录入系统。后来该客户更改了名字，于是员工在系统中更新了相关数据。原先的名字仍旧被系统保留，并作为相关文件的组成部分进行管理和维护。

例 6-2 每年保险单的资产值被自动编入索引，并且该“资产值”字段也是自动更新的。为了在索赔时证明资产的价值，每当更新时，来自“资产值”字段的信息就被移到“先前值”字段。系统要保留过去三年的“先前值”（因为索赔必须在某一事件发生后的三年之内提出），依照业经评注的处置期限表，每年生成新的索赔值。对到期数据的删除，系统要予以记录，包括正式批准的删除。

策略二：将选定的数据元素汇集到一起（可从相同的数据库表中选择数据或从不同的数据库表行里选择数据），并生成固化的、不能更改的、单独的数字对象。此策略可通过生成报告或生成只读的数据库“历史”版本予以实现。

例 6-3 机构使用具有工作流引擎的业务系统来处理贷款申请流程。申请完

毕，系统自动生成一份报告，给出流程的细节，该报告被视为一份文件，随后储存在该机构的电子文件管理系统中。该流程相应的背景信息（以元数据形式）在经 workflow 引擎时被采集，并与文件一起从业务系统导出到电子文件管理系统中。

无论选择何种策略，最重要的，是确保所有核心的文件管理流程均得以运转，这样文件不仅得以生成和管理，同时也能够恰当地予以处置。

例 6-4 一个用于维护客户订单的数据库。依照机构的处置期限表，在订购完成后，订单的详细资料要求保存两年。“查询”程序将在系统中每年运行一次，以鉴别所有之前已完成的、超过两年的订单。这一“查询”结果将由相关人员核对，以确保这些超过两年的已完成订单不与任何悬而未决的订单事宜相关联，并且一旦得到批准，相关的字段即被删除。删除的报告、签字及确认则视为这一流程的证据予以保存。

对流程要谨慎设计，以确保只有与订单相关的字段被删除，而客户的详细资料（要求保留更长时间）则不受影响。

## 二、归档收集方式

电子文件的收集，有在线自动和手动等方式，但归档传递过程必须安全可控，应有相应措施防止在传递过程中丢失电子文件、损失信息和发生非法篡改，以保证电子文件的真实性。

通过业务系统与电子档案管理系统的 Web Service 机制、数据交换等方式进行收集归档的为在线收集。应确认电子文件从业务系统向电子档案管理系统传递过程中不会被非法截获和篡改，并通过系统安全保密设置与运行管理等措施保证归档电子文件在到达电子档案管理系统后不会被非法篡改、丢失、替换等。

### 1. Web Service 在线归档收集

Web service 是使原来各个孤立站点之间的信息能够相互通信、共享而提出的一种接口，是建立可互操作的分布式应用程序的新平台，定义了应用程序如何在 Web 实现互操作。Web Service 所使用的是统一、开放的标准，如 HTTP、XML、SOAP（简单对象访问协议）、WSDL 等。所以，Web Service 可以在任何支持这些标准的环境（Windows、Linux、Unix）中使用。Web Service 具有

平台无关性、普遍性、易于使用等优点，可以沟通不同平台、编程语言和组件模型中的不同类型系统，方便地实现异构系统之间的集成。

SOAP 协议（Simple Object Access Protocol，简单对象访问协议）是一个基于 XML 的通信协议，用于分散和分布式环境下的网络信息交换。在 SOAP 协议下，软件组件或应用程序能够通过标准的 HTTP 协议进行通信。它的设计目标就是简单性和扩展性，这有助于大量异构程序和平台之间的互操作性，从而使存在的应用程序能够被用户广泛访问。

采用 Web Service 方式进行在线归档收集具有以下优点。

- （1）实时性好，在 OA 系统或 PDM 系统中调用 Web Service 方法，可以实时将数据转入到档案系统中。
- （2）通用和扩展性好，双方系统的耦合度低，一方的需求和系统变动不会带来另一方开发工作量的增加。
- （3）安全性好，双方只需沟通传递的数据内容和格式，均不需要了解对方的数据库地址和表结构，不会对对方数据造成潜在危险。

图 6-7 描述了采用 Web Service 方式进行在线归档收集的原理示意图。

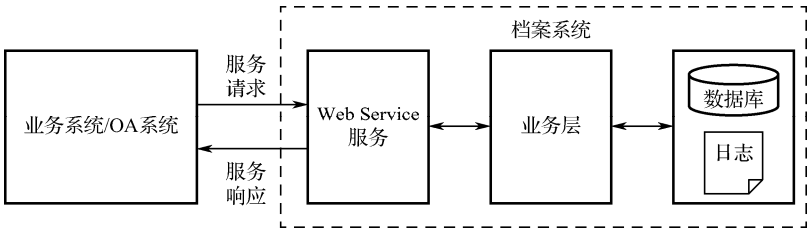


图 6-7 采用 Web Service 方式进行在线归档收集的原理图

如图 6-7 所示，在采用 Web Service 方式进行在线归档收集时，档案系统作为服务提供者，发布 Web Service 服务，其他业务系统（或 OA 系统）作为服务请求者，请求服务。

（1）客户端（其他业务系统或 OA 系统）的主要任务。

① 取得服务端的服务描述文件 WSDL，解析该文件的内容，了解服务端的服务信息，以及调用方式。

② 根据需要，生成恰当的 SOAP 请求消息（指定调用的方法，已经调用的参数），发往服务端。

③ 请求发送前，根据验证的 Schema 文件进行数据格式的验证。

④ 等待服务端返回的 SOAP 回应消息，解析得到返回值。

(2) 服务端（档案系统）的主要任务。

① 编写服务函数代码，生成服务描述（WSDL）文件，以供客户端获取。

② 编写传递的 XML 基础数据集的格式规范，提供 XML 数据示例范本；编写针对该 XML 数据集进行验证的 Schema 文件，用于客户端发送前和服务器端接收后进行格式验证。

③ 接收客户端发来的 SOAP 请求消息，解析其中的方法调用和参数格式。

④ 根据 WSDL 和 WSMIL 的描述，调用相应的对象来完成数据的插入功能。

⑤ 把返回值（成功、失败）放入 SOAP 回应消息，返回给客户端。

⑥ 将调用写入系统日志。

## 2. 数据交换方式在线归档收集

考虑到业务系统和档案管理系统方面的开发成本，业务系统电子文件归档也可采用中间数据交换方式进行归档。数据交换方式包括数据包交换和数据库交换。

采用数据交换方式进行在线归档收集的原理如图 6-8 所示。

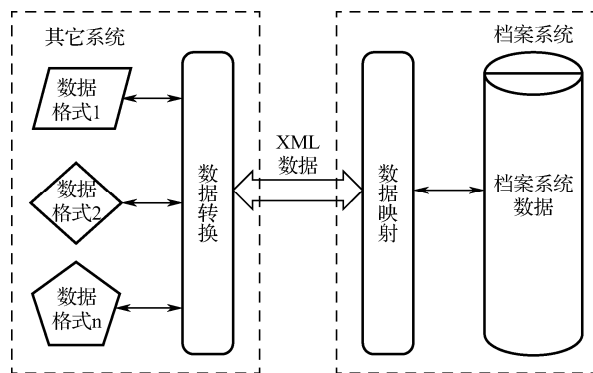


图 6-8 采用数据交换方式进行在线归档收集的原理示意图

如图 6-8 所示，采用数据交换方式进行在线归档收集时，涉及交换数据格式的选择，以及数据转换和数据映射等功能模块。

(1) 数据格式

数据表元数据可以采用统一标准的 XML 格式，数据格式是不同系统间进

行数据交换的接口契约，归档时需要将其他系统的数据转换生成 XML 标准描述的数据结构。

### （2）数据转换

数据转换模块负责在不同格式之间翻译数据，以保证各个系统可以以自己理解的方式接收数据。其他系统的数据需要通过各自的数据转换模块将各类数据格式转换成为统一、规范的标准数据格式（XML），在此基础上再执行后续相关的数据归档处理工作。

### （3）数据映射

数据映射模块解决数据交换过程中不同系统数据结构不一致的问题。通过数据映射功能，将 XML 数据结构映射到档案系统的数据库字段中，完成数据的归档。

在进行中间数据交换时，电子文件可仍旧保留原文件格式，元数据可以采用以下两种方式进行交换。

方式一：将所有元数据以数据表的形式进行存贮和交换。数据表元数据交换格式可选用 XML 和其他对系统依赖度低的格式。

方式二：将审签信息元数据固化为某种格式的电子文件（如 HTML），其余元数据以交换数据表的形式存贮。

存储元数据的电子文件作为电子文件的一部分。对于存储为中间格式的电子文件及元数据，制作成硬拷贝，一式三份，同时按照《信息与文献、文件管理》（GB26162.1）对硬拷贝进行管理。

## 3. 手工归档收集

手工收集是指通过载体移交、拷贝、邮件发送等方式的收集。手工收集应通过现场多人监督等方式确保移交过程中电子文件不被非法篡改。

不管用哪种方式进行归档收集，电子文件归档移交过程中必须将元数据一同移交归档，确保元数据齐全完整。

## 三、归档鉴定

电子文件归档鉴定，是指对电子文件划定保管期限，并对电子文件的真实性、可靠性、完整性和可用性进行检测，检测合格方可归档。

## 1. 划定保管期限

任何电子文件在归档时必须划定保管期限,划定依据应根据机构制定的保管期限表来进行。电子文件归档时,保管期限的划定既可在形成系统电子文件时划定,也可在电子档案管理系统中通过赋予“保管期限”元数据值的方式来划定。

## 2. 真实性、可靠性、完整性和可用性检测

确保电子文件和电子档案真实性、完整性、可靠性和可用性(简称“四性”)是电子文件归档和电子档案管理过程中的一大难题,目前尚无统一的规范和标准。

### (1) 真实性和可靠性检测

主要是检测电子文件和电子档案来源可靠,内容是否被非法更改或非法调换,以及是否完全和正确地表达其所反映的事务、活动或事实。

检测方法如下:

① 收集归档前,采用加密算法对待归档电子文件进行认证,形成认证码。

② 电子档案管理系统接收到归档电子文件后,调用加密算法,将产生的认证码与原认证码进行比对,如果两次认证码不一致,应确定电子文件已被更改。

### (2) 完整性检测

主要是检测电子文件和电子档案信息构成的完整,元数据的完整和特殊软硬件的完整性。

检测方法如下:

① 电子文件信息完整性检测。

将归档电子文件与电子文件应有的构成要素进行比较,凡缺少应有构成要素的电子文件均是不完整电子文件。如一份完整的文书类电子文件一般包括正本及其附件、签发单、修改稿等,如有缺失应认定为不完整。

② 元数据完整性检测。

将收集到的电子文件元数据与元数据定义项进行比较,凡缺少应有元数据元素的,应确定为元数据不完整。

③ 特殊格式电子文件所需的软硬件完整性检测。

对有些需要特殊软硬件支撑才能显示或处理的电子文件,归档时应检查其所依赖的软硬件是否与电子文件一同收集。

### (3) 可用性检测

主要检测电子文件和电子档案的可理解性和可利用性,包括信息的可识别



性、存储系统的可靠性、载体的完好性和兼容性。

① 电子文件和电子档案可识别性检测方法。

第一步：收集不可识别电子文件特征，形成特征库。

第二步：将收集归档的电子文件与不可识别电子文件特征库中的特征进行比对，符合不可识别电子文件特征的应确定为不可识别电子文件。

② 存储系统的可靠性、载体的完好性和兼容性检查。

第一步：采用防病毒系统检查电子文件和电子档案有无病毒；防病毒系统的病毒库应处于最新状态。

第二步：兼容性是指电子文件和电子档案存储格式对特殊软硬件平台的依赖性。兼容性检查包括硬件和软件两个方面。由于当前尚无统一的检查方法，各企业可参照采用信息系统开发中的 Beta 测试法进行检查。

第三步：如果采用离线移交，还应对载体的完好性进行检查，即检查电子文件存储的载体有无损伤，可否正常使用。需检查的离线载体数量较少时，可采用人工检查，需检查离线载体数量较大时，应采用专用检查系统。

## 四、整理

电子文件归档整理包括分类、组成保管单位、保管单位内文件排序、编号等过程。

### 1. 分类

电子档案应以“件”为单位，根据企业制定的分类方案进行分类，并给定类号。

### 2. 组成保管单位

电子文件归档应按照 GB/T 11822、DA/T 22 及其他有关整理的标准、规范组成保管单位。由于电子档案保管期限不受档案盒等物理装具的限制，一个保管单位（如盒）仅是一个逻辑单位，保管单位内文件数据可突破传统档案盒的限制，其数量可以是 100 件、1000 件，甚至 10000 件。但是，尽管电子文件归档的保管单位不受传统物理卷盒的限制，但一个保管期限内电子文件的数量也

不能无限多，应以归档周期为界定组成保管单位，如一周形成的电子文件为一个保管单位，或是一个月形成的电子文件为一个保管单位，也可以一个项目、一次会议形成的电子文件组成一个保管单位。

### 3. 保管单位内文件排序

保管单位内的电子文件应按照一定的逻辑顺序编号进行排序。

### 4. 编制档号

应对完成分类、组成保管单位及排序的电子文件编制档号，档号编制方法参见 GB/T11822、DA/T22 及其他有关的整理标准。

## 五、归档移交

实行业务部门整理立卷制的，当电子文件完成整理并组成保管单位后，其管理权应向档案部门或档案人员移交；对于采用档案部门立卷制的，其管理权可在整理立卷前向档案部门或档案人员移交。移交手续可以通过业务系统、电子档案管理系统设定的功能自动记录，也可以手工填写纸质表格形式进行。但不论何种方式，均需按照 GB/T 18894 有关要求办理移交手续，并保证移交手续不可更改和可查。采用离线方式的，移交过程应确保电子文件不被非法篡改。

依据国家档案局档发[2012]7号《电子档案移交与接收办法》规定如下。

#### 第七条 电子档案移交的基本要求

（一）元数据应当与电子档案一起移交，一般采用基于 XML 的封装方式组织档案数据。

（二）电子档案的文件格式按照国家有关规定执行。

（三）电子档案有相应纸质、缩微制品等载体的，应当在元数据中著录相关信息。

（四）采用技术手段加密的电子档案应当解密后移交，压缩的电子档案应当解压缩后移交；特殊格式的电子档案应当与其读取平台一起移交。

（五）档案移交单位应当将已移交的电子档案在本单位至少保存 5 年。

图 6-9 描述了某银行电子文件归档的封装方案。

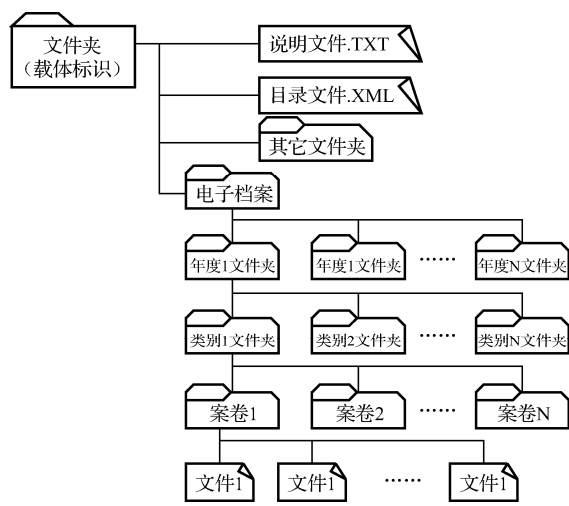


图 6-9 某银行电子文件归档的封装方案

其中，以 XML 格式封装的数据文件示例如图 6-10 所示。

```
<?xml version="1.0" encoding="UTF-8"?>

<dataroot generated="2012-12-14T11:35:46" xsi:noNan
xmlns:xsi="http://www.w3.org/2001/XMLSchema-in>

<案卷>

<ID>7</ID>

<案卷号>111</案卷号>

<案卷题名>2011 年工作会议文件</案卷题名>

<保管期限>永久</保管期限>

<份数>10</份数>

<归档单位>XX 银行</归档单位>

<年度>2011</年度>

</案卷>

<案卷>

<ID>8</ID>

<案卷号>222</案卷号>

<案卷题名>2011 年第十次党代会文件</案卷题名>

<保管期限>永久</保管期限>

<份数>12</份数>

<归档单位>XX 银行</归档单位>

<年度>2011</年度>

</案卷>

</dataroot>
```

图 6-10 某银行电子文件归档中以 XML 格式封装的数据文件示例

图 6-11 显示了其中的说明文件。

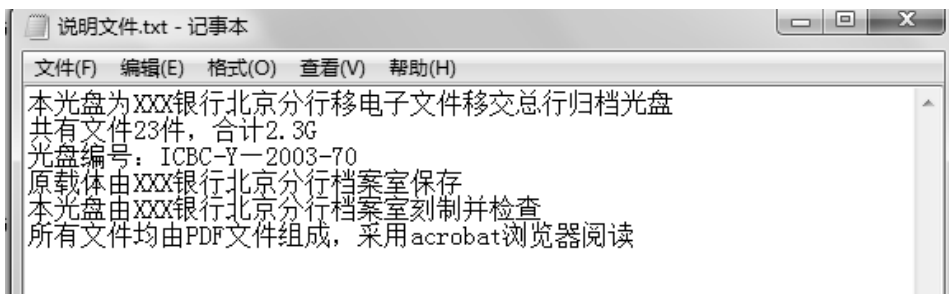


图 6-11 某银行电子文件归档中的说明文件示例

## 六、归档接收

档案部门或档案人员应对业务部门移交的电子文件或电子档案进行清点，并进行真实性、可靠性、完整性和可用性检测，检测合格方可接收。

国家档案局《电子档案移交与接收办法》（档发[2012]7 号）中，与电子文件归档接收相关的主要条文如下。

第十四条 电子档案接收的主要流程是：检验电子档案数据、办理交接手续、接收电子档案数据、著录保存交接信息、迁移和转换电子档案数据、存储电子档案数据等步骤。

第十六条 电子档案检验合格后办理交接手续，填写《电子档案移交与接收登记表》（见图 6-12），由交接双方签字、盖章，各自留存一份；《电子档案移交与接收登记表》可采用电子形式并以电子签名方式予以确认。

图 6-12 为《电子档案移交与接收登记表》样例。

电子档案移交与接收登记表

交接工作名称			
内容描述			
移交电子档案数量		移交数据量	
载体起止顺序号		移交载体 类型、规格	
检验内容	单位名称		
	移交单位:		接收单位:
准确性检验			
完整性检验			
可用性检验			
安全性检验			
载体外观检验			
填表人（签名）	年 月 日		年 月 日
审核人（签名）	年 月 日		年 月 日
单位（印章）	年 月 日		年 月 日

图 6-12 《电子档案移交与接收登记表》样例

第四节 电子档案存储、保管、利用、处置

一、电子档案存储

1. 电子档案在线存储

电子档案的在线存储方式，应在电子档案管理系统设计开发时进行规划。选择的存储方式应有利于保证电子档案真实、完整、长期可用，有利于保证电子档案的安全，成本最小。

## 2. 电子档案离线存储

有条件的企业应对电子档案进行以单份文件（含元数据）为存储单元的离线存储，以进一步降低电子档案长期保存风险。

重要的电子档案应进行一式三套离线存储，载体应具有较好的耐久性，载体选择标准依次为一次性写光盘、磁带、可擦写光盘、硬磁盘等。离线载体应按要求进行标识并按照磁性载体档案进行保管。三套载体最好分开保存，有条件的单位应将一套置于距离 300 公里以上、不在同一流域的地点保存，或不同的建筑物保管。

离线存储载体管理按照 DA/T38《电子文件归档光盘技术要求和应用规范》和 DA/T15《磁性载体档案管理与保护规范》进行。

离线存储时，应按《电子文件移交与接收办法》（档发[2012]7 号）和离线存储载体容量进行信息组织，不能用运维备份的文件组织方式来进行离线存储，更不能用系统备份文件代替离线存储文件。

对磁性载体中的电子档案，每满 2 年，光盘每满 4 年进行一次抽样机读检验，抽样率不低于 10%。如发现问题应及时采取相应措施保证载体上的电子档案长期可读。

对磁性载体上的电子档案，应每 4 年转存一次，原载体同时保留时间不少于 4 年。

## 二、电子档案保管

电子档案的保管，是指采用一定的方法，确保电子档案得到长期真实维护的过程和操作。这些过程和操作包括在线、近线和离线存储环境管理，迁移、技术典藏、封装、环境虚拟、仿真、检测、介质选择等。

### 1. 离线环境要求

所有电子档案的存储环境，必须能够防止直接的太阳光和紫外线照射，免受化学或者特殊大气污染，避免温度和湿度的频繁波动。

保管磁盘介质的环境不能有电子设备或者其他会产生磁场的设备；同时，周边环境的磁场强度也不应该超过 4000A/m。

建议存储环境应该根据数字资源保存的期限，不同的存储期限采取不同的环境条件。而档案管理部门对存储环境的投入成本决定存储介质的保存期限。

(1) 短期存储

表 6-1 描述了短期存储环境条件表。

表 6-1 短期存储环境条件表

介 质	温 度	相 对 湿 度
Flexible Magnetic Disks	10℃～51.5℃	20%～80%
Digital Audio Tape(DAT)	3℃～45℃	20%～80%
Digital Linear Tape(DLT)	16℃～32℃	20%～80%
Uhrium Linear Tape Open(LTO)	16℃～32℃	10%～80%
Other Magnetic Tape Cartridges	10℃～45"12	20%～80%
CD -ROM/R/RW	10℃～50℃	10%～80%
DVD-ROM/R/+R/RAM/Rw/+RW	10℃～50"12	10%～80%
固体/电子晶体介质 Solid State Media	Media 10℃～50℃	20%～80%
混合馆藏	16℃～32℃	20%～80%

(2) 长期存储

表 6-2 描述了长期存储环境条件表。

表 6-2 长期存储环境条件表

介 质	温 度	相 对 湿 度
Flexible Magnetic Disks	18℃～22℃	35%～45%
Digital Audio Tape(DAT)	5℃～32℃	20%～60%
Digital Linear Tape(DLT)	18℃～26℃	40%～60%
Uhrium Linear Tape Open(LTO)	16℃～32℃	20%～80%
Other Magnetic Tape Cartridges	18℃～22℃	35%～45%
CDROM/R/RW	18℃～22℃	35%～45%
DVD-ROM/R/+R/RAM/Rw/+RW	18℃～22℃	35%～45%
固体/电子晶体介质 Solid State	18℃～22℃	35%～45%
混合馆藏	18℃～22℃	35%～45%

## 2. 迁移

迁移被认为是数字信息长期保存的必须和恰当的选择,也是目前比较成熟和频繁使用的方法。通过这种方式获得对软硬件环境的相对独立性,保证电子文件在新的环境中能够被识别。档案管理部门应每年对电子档案的可读取性进行评估,形成评估报告;如存在因系统软硬件或其他技术升级、变动出现电子档案不可读取的风险,应对电子档案进行迁移。

国家标准《电子文件归档与管理规范》指出,迁移是指将源系统中的电子文件向目的系统进行转移存储的方法与过程。延长电子文件技术寿命的迁移,则是指将电子文件从过时的软硬件平台向新的软硬件平台的技术性转移,这种转移通常伴随着存储格式的改变或版本的升级。

迁移依靠运行于新技术平台上的应用软件来实现。通常,升级后的应用软件具有一定的向下兼容性,即能够打开前几代版本生成的档案文件。这些旧的档案文件在升级版系统中打开并重新存储后,便完成了电子文件向新技术平台的迁移。某些情况下,由于电子文件信息的存储格式已遭淘汰,还需要采用适当的转换工具,将使用旧格式存储的电子文件转换成新的通用或标准格式的档案文件。这种技术性转移已经成为对抗技术平台升级而导致电子文件失效的较为经济、也更为常用的对策。

事物总是有两面性的,迁移策略的实施也面临着诸多风险。

一方面,迁移过程中电子文件信息失真情况比较严重,特别是涉及文件格式迁移时,常常会丢失电子文件原有的某些功能或信息,如图表展示、索引、图层、链接等。因此,迁移技术较适用于那些与软件无关的中间性格式文本档案文件或简单、通用的平面档案文件,而不适合结构复杂的电子文件,如多媒体作品、交互性档案文件、联机对话等动态数据。连续不断的迁移,失真将会逐步累积,最后导致电子文件面目全非。

另一方面,随着技术环境的不断变化,人们必须不断选择时机进行迁移,而迁移时机的把握十分困难,迁移时间过晚,可能导致电子文件再也无法识读,有效性彻底丧失,而迁移时间过早,迁移过于频繁,不仅会造成巨大的资源浪费,而且会加大电子文件信息失真的概率。要准确把握迁移时间,必须跟踪技术发展状况,及时检测技术更新对电子文件有效性的影响。因此,运用迁移方法维护电子文件长期的有效性需要对技术本身不断完善,并研究制订迁移对象,筛选和确定迁移时间。



电子档案迁移前应进行迁移可行性评估，包括目标载体、系统、格式的可持续性评估、成本评估等，并保证迁移过程电子文件真实性，过程可控，防止迁移过程中电子文件信息丢失、非法篡改。迁移后原载体保留时间不得低于三年。

迁移包括以下具体步骤。

#### （1）迁移规划

首先制定迁移规划，对数据或系统进行迁移时，应先对新环境软硬件进行选型，选型指标主要包括适配时间长，产品成熟，性能稳定，运行效率高等。通过新环境选型，确定最优的软硬件组合，然后对数据进行迁移，迁移过程中，应对与环境不兼容的数据进行格式转换，进行合理迁移。数据迁移完成后，对其进行测试、验证，以确保在新环境中可以正常使用。

#### （2）迁移方案

根据迁移规划，确定迁移所涉及的软硬件和网络等内容，制定迁移计划时间表，按照迁移的实际情况，安排迁移系统或数据及网络情况等优先级。一般情况下可先迁移数据，然后再进行对应系统的迁移，最后进行相关网络的配置，并且对此环境确定备份方案、迁移步骤、迁移周期和参与人员等。

#### （3）迁移评估

在对应用系统进行迁移之前，首先应明确新环境下所有系统的网络环境、资源利用率、服务器和各系统需求、数据对现有资源的利用率、文本格式、音视频格式之间的差异后，对迁移进行计划，从而保证各应用系统从旧环境迁移到新环境的过程中风险最小。

#### （4）迁移实施

在数据迁移之前，应明确迁移起止时间，明确团队成员的角色。例如，针对迁移中央档案馆真实常用数据 2T 以上的要求，应合理分配人员，以确保在最短的时间内完成数据迁移。然后执行备份方案，对需要迁移的数据及系统进行备份，执行相关的迁移步骤。

### 3. 技术典藏

面临技术过时带来的读取危机，人们想到的最原始的办法就是将运行电子文件的软件、硬件平台连同电子文件一起保存下来，以备今后能够读取那些过时的电子文件信息内容，延长其技术寿命。

在其原生环境中识读，无需迁移环境或转换格式，电子文件信息不会失真，信息处理的各种功能也都能得到完整保留。如果这样能够得以顺利实施，这将是一种对付技术淘汰，维护电子文件长期有效的“最理想”的方法。

但事实上这种方法所具有的致命缺点却证明了它只是一种“理想化”的方法，并不具有可行性。不难想象，试图将不同时期、不同技术背景下的电子文件和软硬件平台全都完整地保留下来决非易事，需要建一个庞大的软硬件“博物馆”。

(1) 一方面它必须收存运行电子文件的所有硬件设备，而这些硬件设备的保存受到物理生命的限制，自身会日渐损坏，不可能长期存留，而人们也不可能将该硬件设备的整条生产线保存下来。

(2) 另一方面它要收存所有设备的驱动程序、各种版本的操作系统、数据库系统和其他平台软件，尤其是要保留生成专门格式电子文件品种繁多的应用软件，这些软件与硬件之间、系统软件与应用软件之间互相支撑、相互关联，缺一不可。

因此，这个“博物馆”不仅耗资巨大，而且维护起来极其困难，仅凭档案部门的技术和资金力量是难以承担和维持的。即使这个软硬件“博物馆”建立起来而且维护良好，电子文件信息的读取利用也会受到空间的制约，软硬件设备都被保存在“博物馆”内，地理位置不易改变，远程的利用需求难以满足。

在诸多的存储介质中，微缩胶片已经逐渐被数码摄影、PDF 等取代，因为阅读微缩胶片必须借用特制仪器—柜式胶卷机。它在购买、维修、贮藏上的成本很贵。但为了防止已电子化的档案因储存媒体及档案格式迅速变迁，而无法被未来的电脑读取及处理，除了电子化以外，仍继续使用微缩胶片保存重要的档案，这就需要在保存胶卷和微缩胶片的同时也要将阅读设备（柜式胶卷机）一并保存。

#### 4. 封装技术

在程序设计技术中，封装技术是指隐藏对象的属性和实现细节，仅对外公开接口，控制在程序中的读和修改的访问级别；将抽象得到的数据和行为（或功能）相结合，形成一个有机的整体，也就是将数据与操作数据的源代码进行有机结合，形成“类”，其中数据和函数都是类的成员。

封装是把过程和数据包围起来，对数据的访问只能通过已定义的接口。面

面向对象计算始于这个基本概念，即现实世界可以描绘成一系列完全自治、封装的对象，这些对象通过一个受保护的接口访问其他对象。封装是一种信息隐藏技术，在 java 中通过关键字 `private` 实现封装。属于面向对象编程语言中的概念，属于软件范畴。

封装技术是指实体文件与元数据封装在一起。

5. 环境虚拟

环境虚拟也称计算机虚拟化技术。通常是指计算元件在虚拟的基础上，而不是真实的基础上运行。虚拟化技术可以扩大硬件的容量，简化软件的重新配置过程。CPU 的虚拟化技术可以单 CPU 模拟多 CPU 并行，允许一个平台同时运行多个操作系统，并且应用程序都可以在相互独立的空间内运行而互不影响，从而显著提高计算机的工作效率。

6. 仿真技术

仿真技术在档案领域的使用有两种形式。

一种是利用扫描、数码相机、图像输出等现代化手段对原档案进行仿真复制。此种方式多用在传统纸质档案的仿真处理，在一般的利用过程中，不需要原版的情况下，可以使用仿真件，最大限度地降低原件的使用率，起到了保护原件的作用。

另一种是在新的软硬件平台上，通过运行某种仿真软件来“虚拟”旧的技术平台，从而使旧技术平台上生成的数字档案能够在新的软硬件环境中识读和处理，这种方式会用在电子档案的仿真处理上。仿真软件作为识读、处理旧档案文件的工具，其功能是实现旧技术平台向新技术平台的“影射”，如图 6-13 所示。

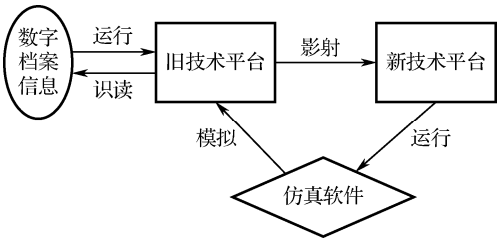


图 6-13 利用仿真软件识读和处理档案文件

显然,从理论上来说,仿真有一定优势。一方面,通过仿真软件可以很方便地运行旧的电子文件,无需实际配置旧的技术环境,而且针对某一类电子文件而研制。仿真软件可再用于存取同一时期相同类型的其他档案文件;另一方面,通过仿真软件所读取的旧电子文件能以其原始面貌示人,相关的信息处理功能也能得以保留。

但是,仿真技术方法也存在人们无法回避的实际缺陷。一方面,仿真软件的开发难度很大,代价很高,每一个仿真软件只适用于特定的新旧平台,兼容性较差。随着软硬件的不断升级,仿真软件也必须随之升级,以适用于新的技术平台。而对仿真软件升级维护的代价相当昂贵,开发仿真软件的费用可能将超过反复迁移费用的总和。另一方面,仿真软件不可能绝对真实和完整地模拟出电子文件当初运行的技术环境,其仿真度受到质疑,在仿真软件上识读旧的电子文件可能会丧失部分功能,出现信息失真。因此,这种方法仅作为延缓技术淘汰的阶段性的方法,用于部分对特定软硬件具有很强的依赖性或者在新旧技术环境之间进行迁移有很大困难的电子文件的保护和读取。

电子会计档案包括电子凭证、电子账簿、电子报表、其他电子会计核算资料等。这些数据在电子化过程中可能被保存在磁带等磁性介质上,这些利用磁带存储的电子档案在存储时不能只存储在磁带本身,因为磁带本身的信息并不能直接被人们所识别,需要通过相应的阅读设备才能读取存储在其内部的信息。所以在保存磁带的同时,磁带机等磁带阅读设备一并保存。但是磁带对保存环境的温度、湿度、磁场等都有较严格要求,并且必须每隔10年复制一次,其性质不利于长期保存,并且磁带阅读设备也有相应的使用年限,一旦阅读设备失效,那么磁带上记录的档案信息也失去保存意义。在这种情况下,若要长期保存这部分电子会计文档,需要利用仿真技术虚拟磁带机运行环境,通过不断在新的技术平台上虚拟旧的数字环境来实现旧数据的长期保存。

另外,在档案管理系统中使用仿真技术,在法律层面上,现在仍然具有不确定性。所以,目前仿真件在简单的档案借阅利用场景下可以使用,一旦涉及法律凭证等方面还需要使用原件。

## 7. 介质选择

选择合适的存取介质,主要取决于IT提供商、文件档案信息存取访问级别以及投入成本。限制条件主要是介质的长期可获得性。最好是该介质能防止档

案信息内容和质量的退化，换句话说，越稳定的介质对档案数字资源的长期保存越有益。

尽管光盘 WROM 存储介质从理论上讲是最可靠的，但是因存储容量和存储成本因素，采用可擦写式的磁盘介质也是可取的。

由于软件格式不断更新变化，档案管理部门有必要考虑为档案利用者提供存取访问策略。重要的是，如果档案管理部门决定采用阅读器技术为用户提供一定的档案信息，那就要保证该阅读器技术能够存取访问所有格式的档案文件。如果不采用阅读器技术，应该把相关文件迁移为一定的格式，使之能够被查看、阅读。

各种类型的存储介质都有一定的生命周期，同时，该周期会受环境因素（热度、温度、湿度、地磁场）的影响。为了防止档案数字资源的质量下降，必须在一定条件基础上，把这些存储在介质中的档案信息重写到新的介质中，而重写的频率必须考虑介质本身的性能。介质稳定性检查的时间间隔不能低于 5 年。

Adrian Brown 在《数字档案分析家：为长期保存选择存储介质》一文，对目前市面上较流行的六种存储介质进行评分，结果见表 6-3。

表 6-3 长期保存介质时间表

介质	CnR	DVDtR	Z1D Dlsk	3. 5in 磁盘	DLT	DAT
长期性	3	3	1	1	2	1
存储容量	2	2	1	1	3	3
健壮性	2	2	1	1	3	3
过时	3	2	2	3	2	2
成本	3	2	1	1	3	3
敏感程度	3	3	1	1	3	2
总分	16	14	7	8	16	14

国标 GB/T 18994 规定，推荐采用的载体，按优先顺序依次为：只读光盘、一次写光盘、磁带、可擦写光盘、硬磁盘等。不允许用软磁盘作为归档数字文件长期保存的载体。

### 三、电子档案利用

企业应制定详细的电子档案利用权限的规定，利用权限应在电子档案管理系统中实施，并经过确认。当超权利用时，必须进行审批，并保证利用过程中电子档案不被非法篡改。

电子档案可根据授权，通过档案信息系统在线或离线提供利用，但不管何种利用方式均应通过日志或其他方式记录利用过程，记录信息包括利用人、利用方式、电子档案档号、文件号、文件名、利用时间等。利用过程信息应作为电子档案元数据的一部分进行保存。

### 四、电子档案处置

电子档案的处置包括定期鉴定、销毁及进馆。

#### 1. 定期鉴定

定期鉴定是指机构对保管期限已到的定期保管的电子档案进行鉴定，以确定其处置行为的过程。

鉴定方法可根据不同类别档案的管理要求进行。

#### 2. 销毁

电子档案的销毁应先登记，编制清册，并按有关规定履行审批手续后，再将电子文件进行物理删除；删除时至少需要两人监督，销毁清单及记录需打印纸质档案保存。涉密电子文件的销毁，应当按照国家保密法律法规的规定处理。

#### 3. 移交进馆

企业应按国家有关规定要求，将应进馆的电子档案及时移交进馆，包括纳入进馆范围的电子档案及其元数据，移交方法参见《电子档案移交与接收办法》。

# 第七章 电子档案管理系统软件 分析与设计

## 第一节 电子档案管理系统软件生命周期

### 一、软件开发周期阶段

软件开发周期，是为每个软件开发项目而进行管理活动的一个有序集合。软件开发周期可以在不同阶段的粒度上提出，在粗粒度层面上，软件开发周期可以包含分析、设计、实施三个阶段。分析阶段专注于系统需求，需求在这里被确定，系统功能和数据模型被开发并集成在一起，非功能需求以及其他系统约束也被捕捉出来。设计阶段可以划分为几个子阶段，分别为体系设计和详细设计，这个阶段需要解释用来集成用户界面和数据库对象的客户机/服务器程序的设计。需要提出影响系统的可理解性、可维护性和可扩展性等各种设计方面的问题并文档化。实现阶段则采用增量迭代的方式通过编码设计，将设计方案转换成可执行的程序。

在细粒度层面，软件开发周期划分为需求确定、需求规格说明书编制、体系结构设计、详细设计、实现、集成和维护。

### 二、系统生命周期阶段

电子档案管理系统大多通过购买方式取得，本书采用粗粒度划分，将电子档案管理系统开发划分为分析、设计、实施三个阶段。在电子档案管理系统实际采购工作中，分析、设计、实施三个阶段的工作都由系统开发商来完成，档案工作人员只是参与其中。但是，由于三个阶段工作内容不同，需求阶段要求档案人员作为系统用户要更多地参与其中。一般来说，三个阶段档案人员与系

统开发商分工见表 7-1。

表 7-1 系统开发任务分工表

生命周期阶段	工 作 内 容	用 户	软件开发商
需求分析阶段	需求确定和形成需求规格说明	参与	负责
设计阶段	详细设计	参与	负责
实施阶段	软件代码及测试	参与	负责

1. 分析阶段

电子档案管理系统分析阶段，主要工作是需求确定和形成需求规格说明。

(1) 需求分析

即电子档案管理系统服务或约束的陈述，包括为系统提供功能和其他需求的陈述性定义，系统被期望的服务（服务陈述）和系统要服从的约束。服务陈述包括系统的范围、必要业务功能（功能需求）和要求的数据结构（数据需求）。约束性陈述包括系统的“外观和感觉”、性能、安全性等。

(2) 需求说明

即开发者使用一种特定的方法（如 UML）对需求进行建模，形成一个由图形模型和工具生成的表格文档。在电子档案管理系统开发时，需求说明工作一般由系统开发人员完成，但档案人员需要对其进行确定。

2. 设计阶段

电子档案管理系统设计阶段，可划分为概要设计和详细设计两个子阶段。概要设计给出软件结构、全局数据结构、数据库结构和接口；详细设计给出各模块的具体实现算法和彼此的调用关系。这是一个把语言描述转换为更精确的、结构化的过程描述。

3. 实施阶段

即电子档案管理系统实现阶段，是把设计阶段的过程描述用某种计算机语言代码来表示（也即编码）。

由于本书的读者主要为档案部门人员，关心的是需求设计，因此，本书主要阐述电子档案管理系统的需求分析。



## 第二节 档案管理信息系统需求分析

需求分析是系统开发中更不需要技术的一个阶段，但是，如果这个阶段不完全进行，其结果将会使其他阶段更糟糕。由于没让软件开发商正确地捕获需求，使软件开发商忽略或错误地理解需求，软件用户将为此付出不可承受的代价。在实际工作中经常有企业档案部门要求软件公司补充开发，从而引发商务纠纷，这些现象的出现都与当初不进行充分的需求分析有关。需求分析是档案信息化建设过程中的重要工作，特别在购置电子档案管理系统软件时，如果没有进行系统而全面的需求分析，往往在实施过程中出现功能或业务遗漏或业务流程错误。

需求分析包括需求获取、需求陈述和形成需求分析报告等。

### 一、需求获取方法

#### 1. 原型法

原型法在需求获取中是最常见的方法。要使用原型法首先需要构造软件原型，原型是一个演示系统，它是解决方案的一个“快速而粗糙”的工作模型，它呈现图形用户界面，并模拟用户各种事件行为。系统原型是从用户那里抽取，其他方法难以获取需求的一种有效方式。在增加新业务功能时，常常用这种方法，对用户与开发人员之间沟通问题时具有重要作用。但现在进行电子档案管理系统开发时，常常用原有系统或开发商现有的系统作为原型。

原型法也用于对现行系统的详细调查。因为现行系统的详细调查，是通过各种方式对现行系统做详细、充分和全面的调查，弄清现行系统的边界、组织机构、人员分工、业务流程、各种计划、单据和报表格式、处理过程、企业资源及约束情况等，使系统开发人员对现行系统有一个比较深深刻的认识，为新系统开发做好原始资料的准备工作，这是需求分析中一项必要的工作。

#### 2. 面谈法

有时电子档案管理系统用户对需求只有一个模糊的认识，他们也可能不愿

意合作或不能够表达他们的需求，这时用面谈法比较有效。面谈法有结构化面谈和非结构化面谈两种形式。结构化面谈需要提前准备，有一个明确的日程，并且许多问题都是预先确定的，有一些问题可以没有明确的答案。非结构化面谈就好像非正式会议，没有预定的问题或预定的目的，其目的是鼓励用户阐述其想法，并且在这个过程中提出业务分析人员没有想到的和没能提出的相关需求。

面谈法必须有某个出发点和讨论话题，这也可以是一篇短小的书面文档，或是发给面谈对象的电子邮件。

面谈法是软件开发人员与用户确定需求的重要方式，尤其在向企业高层管理者、非档案人员及缺乏信息技术知识的档案人员收集需求时采用。

### 3. 观察法

观察法用于面谈法和原型法都无法获取需求的情况，即分析人员通过现场观察，掌握业务进行的方法。如系统分析人员往往无法从面谈法中理解档案整理的业务需求，通过观察档案人员手工现场组卷，或观察档案人员现场操作，就是一种最有效的方法。

### 4. 文档和软件系统研究

这是获取需求和相关领域知识的最好方法。

在电子档案管理系统开发中，要研究的用户文档包括：业务表格、工作过程、职位描述、档案管理标准规范、组织图档、各种报表等。

要研究用户的系统表格，主要是旧系统表格，包括：系统操作手册、用户文档、技术文档、系统分析和设计模型等。

还经常用到相关领域知识，可通过档案期刊或参考书来获取。可将档案行业标准规范提供给需求分析人员，如在确定元数据需求时，常常将《文书类电子文件元数据方案》（DA/47）提供给分析人员。

二、需求陈述方法

需求分析的内容包括功能需求、性能需求、领域需求和其他需求。这些需求要通过某种描述性语言将它固定下来，这就是需求陈述。

1. 功能需求陈述

功能需求主要说明系统应做什么，这是用户最直观也是最主要的需求，如系统的输入输出、系统能完成的功能，以及其他相关处理等。

如“系统支持对接收的归档电子文件进行‘四性’检测，检测不合格的不接收”，具体如图 7-1 所示。

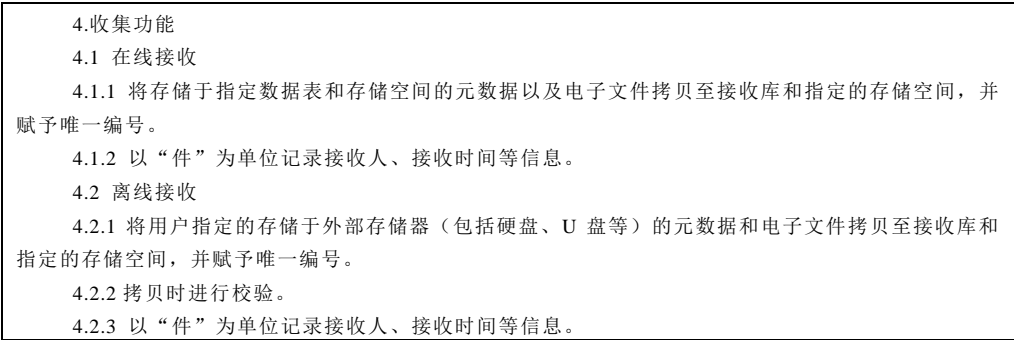


图 7-1 功能性需求陈述

2. 约束性需求陈述

约束性需求陈述主要从各个角度对系统起约束和限制作用的陈述。如系统的响应时间、存储效率、报表规格和界面样式等，主要是系统在空间、时间上的约束，一般称为系统性能。但在进行需求分析时，往往会把系统性能独立出来，从标准规范性考虑，系统性能指标有可能在功能描述时一同描述。约束性需求陈述如图 7-2 所示。

3. 领域需求

领域需求即专业需求。即要求系统功能和数据需要特别实现、具有特殊性

的需求。如电子档案管理系统元数据项必须符合《文书类电子文件元数据方案》(DA/47),这就是一种领域需求。

■性能需求

打开页面或提交事件响应时间 1-3 秒;

综合查询响应时间小于 5 秒;

全文检索响应时间小于 30 秒;

系统硬件设备的 CPU 忙时利用率平均不超过 75%,内存忙时利用率平均不超过 75%;

对于大数据量的统计,应能保证在夜间(4 小时内)完成;

系统消息处理延迟(从系统收到请求消息到返回回答的时间间隔)不大于 500 毫秒(99.999%概率)。

■安全性要求

任何用户访问系统资源,必须得到系统的身份认证以及身份标识,如用户的数据证书、用户号码、密码。当用户信息与确认信息一致时,才能获准访问系统。在本系统中,对操作系统,数据库系统和应用系统都有相应的用户和权限的设置。

对系统资源,包括程序、数据文件、数据库等,根据其特性定义其保护等级;对不同的用户,规定不同的访问资源权限,系统根据用户权限,授予其不同等级的系统资源的权限。

在系统中所保护的资源进行任何存取操作,都做相应的记录,形成日志存档,完成基本的审计功能。

在系统中对在网络中传输的信息必须经过高强度的加密处理来保证数据的安全性。

系统应可防止网络安全或漏洞入侵,安全必备的服务器防病毒软件,且漏洞补丁的部署不影响系统的使用。

■可用性要求

平台要考虑 7X24 小时运行,充分考虑系统容错机制,每年非计划内宕机不得超过 8 小时。

系统支持双机热备,在集群方式下为 N+1 备份,在非集群方式下为 1+1 备份。

从数据库备份到备份系统的时间:每天的备份时间不多于 1 小时,而且不影响对数据库的少量查询。

从备份系统拷贝到数据库里的时间:数据恢复时长不多于 1 小时。

系统平均无故障时间不小于 26280 小时(3 年)。

平台支持负载均衡,系统切换时间小于 10 秒。平台在切换过程中应保证已经处理过的业务正常运行(满足容灾 7 级要求,国际标准 SHARE78)。

系统平均无故障率不低于 99.999%。

■易用性要求

系统对用户界面应保证交互性和美观性,保证用户体验和使用效率。具体体现在以下几个方面。

易理解:单凭凭观察,用户可在几秒内就知道系统是做什么的,以及知道如何采取行动。

易操作:用户初次接触网站时,可以很轻松地完成任务,可以不用帮助文档就能使用网站。

易学习:通过简单的帮助文件,用户可顺利解决遇到的问题。

图 7-2 约束性需求陈述

## 4. 需求管理

### (1) 需求识别与分类

需求用自然语言描述,通常有几十甚至上百条陈述语句组成。要恰当管理好这样大量的需求,首先要对需求进行分类,然后用层次编号法进行编号,对

每条需求赋予唯一编号。

### （2）需求层次

进行了编号的需求，可按父子关系建立层次化结构，父级需求由子级需求构成，子级需求是父级需求的子需求。

### （3）需求的协商

在需求识别和编号中，往往会发现从用户处获取的需求可能是重叠和矛盾的，有些需求可能是二义或不可实现的，而其他需求又可能没有发现，因此，需求需要协商和验证。

## 5. 需求分析与手工管理关系的处理

在需求分析中，用户常常提出一些系统无法实现或实现起来不经济的需求，特别是在整理环节。这时，千万不能因为系统无法实现而丢弃该需求，也不能用大量的编程量去满足一些细微的只有手工才能实现的需求。可用的方法是，在手工需求和系统可实现的功能之间找到一个兼顾两方的需求。即该需求既可解决手工操作转换为系统操作，又不致于花费太多开发力量。

## 6. 需求分析中要注意的问题

有几种原因使需求分析变得困难：一是用户说不清楚需求；二是需求自身经常变动；三是分析人员或用户理解有误。

### （1）客户说不清楚需求

有些用户对需求只有朦胧的感觉，说不清楚具体的需求。例如，全国各地的很多机构在搞网络建设，这些单位的人员大多不清楚计算机网络有什么用，反而要软件系统分析人员替他们设想需求。这类工程的需求比较主观，以致产生很多浪费现象。

有些用户心里非常清楚想要什么，但却说不明白，可能不能引起系统分析人员的注意。比如买鞋子，非常了解自己的脚，但没法说清楚脚的大小和形状，只能拿鞋子去试，试穿时感觉舒服才会买。

如果用户自己就懂软件开发，能把需求说得清清楚楚，这样的需求分析将会非常轻松。如果用户不擅长软件，就必须信任软件分析人员。因为，分析人员可以引导用户，先阐述常规的需求，再由用户否定不需要的，最终确定用户真正的需求。最怕的就是“不懂装懂”或者“半懂充当内行”的用户，他们会

提出不切实际的需求。

#### (2) 需求自身经常变动的应对方法

① 系统分析时应尽可能分析清楚哪些是稳定的需求，哪些是易变的需求。以便在进行系统设计时，将软件的核心放到稳定的需求上，以减少反复。

② 在合同中一定要说清楚“做什么”和“不做什么”。如果合同含含糊糊，日后扯皮的事情就多。

#### (3) 分析人员和顾客理解有误

软件系统分析人员不可能都是全才，用户表达的需求，不同的分析人员可能有不同的理解。如果分析人员理解错了，可能会导致开发人员白干活，吃力不讨好。就如写作文跑题，不管作文写得有多长，总是零分。所以分析人员写好需求说明书后，需要请各方面用户代表进行验证。如果问题很复杂，双方都不太明白，就有必要请开发人员快速构造软件的原型，双方再次论证需求说明书是否正确。

由于档案用户大多不是信息技术专业，会提出一些无法实现的需求，从而无法接受软件系统分析人员的建议或答复。

## 三、系统分析报告

系统分析报告是固定需求分析成果的文档，也是需求分析成果的体现。

### 1. 系统分析报告的组成

一份完整的系统分析报告应该包括以下内容。

#### (1) 系统概述

- ① 目标系统的名称、目标和主要功能。
- ② 背景、系统的用户、开发者以及本系统与其他系统或机构的关系和联系。
- ③ 参考资料和专门术语说明。

#### (2) 现行系统概述

① 现行系统现状调查说明。通过现行系统的组织结构图、数据流图和概况表等图表及说明，说明现行系统的目标、规模、主要功能、组织机构、业务流程、数据存储和数据流以及存在的薄弱环节。


② 各级组织需求说明。用户要求以及现行系统主要存在的问题等。

## (3) 目标系统需求

① 系统目标、特点、假定和约束。

② 需求规定。

组织架构, 权限需求, 任务处理流程需求, 详细功能需求 (包括功能解释, 数据描述、系统管理、运行环境规定、开发技术约定等)。图 7-3 是分析报告的内容要点, 供参考。但不同的软件开发人员形成的需求分析报告会有细微差别。



2.1 目标  
2.2 系统的特点  
2.3 假定和约束  
3. 需求规定  
4.1 组织架构  
4.2 权限需求  
4.3 任务处理流程需求  
4.3.1 系统用例图  
4.4 详细功能需求  
4.4.1 功能解释  
4.4.1.1 收集功能  
4.4.1.2 整理功能  
4.4.1.3 保管功能  
4.4.1.4 统计功能  
4.4.1.5 鉴定销毁功能  
4.4.2 数据描述  
4.4.3 系统管理  
4.4.3.1 部门管理  
4.4.3.2 员工管理  
4.4.3.3 用户管理  
4.4.3.4 权限管理  
5. 运行环境规定  
5.1 设备  
5.1.1 客户端  
内存: 512M 以上  
CPU: 主频 1.6G 以上。  
操作系统: Windows Xp。  
浏览器: Internet Explorer 6.0 以上。  
5.1.2 服务器端  
内存: 2G 以上。  
CPU: 主频 2G 以上。  
操作系统: Windows Server 2003。  
硬盘: 120G 以上。  
5.2 支撑软件  
后台数据库采用 MySql。  
Web 服务器采用 Tomcat。  
6 开发技术约定  
系统采用 B/S 架构, 即 Browser/Server 架构。  
开发语言采用 Jsp、Java 语言。  
系统功能页面的分辨率为 1024×768。

图 7-3 系统分析报告要点

(4) 系统设计与实施初步计划

① 工作任务分解。根据资源及其他条件确定各子系统开发的先后次序，在此基础上分解工作任务，落实到具体组织或个人。

② 时间进度安排。

③ 预算，对费用的进一步估计。

(5) 领导审批意见

## 2. 系统分析报告的作用

① 描述目标系统的逻辑模型，作为软件系统服务商进行系统修改设计与实施的基础。

② 作为用户和软件系统服务商之间的协议或合同的一部分，为双方的交流和监督提供基础。

③ 作为目标系统验收和评价的依据。

因此，系统分析报告是系统购买或开发过程中的一份重要文档，要求该文档必须完整、一致、精确，且简明易懂，易于维护。

## 第三节 电子档案管理系统需求陈述

本节根据前述内容，提出企业数字档案馆需求的一般陈述，供读者参考。

### 一、收集功能

#### 1. 电子文件手工登记

将手工输入的电子文件赋予唯一编号。所输入的元数据项在满足国家有关标准要求的前提下，可根据需要增减；具有元数据校验功能；可对电子文件元数据根据权限进行更改、删除、检索、全文挂接等操作。



## 2. 电子文件和电子档案在线接收登记

具有从业务系统接收电子文件及其元数据的接口；对所接收的电子文件赋予唯一编号。

从其他电子档案管理系统接收电子档案的功能，功能符合《电子档案移交与接收办法》有关要求。对所接收的电子档案以件为单位给予唯一编号。

## 3. 电子文件和电子档案离线接收

从离线载体批量导入电子文件、电子档案和其他数字资源，支持常见的XLS、DBF、MDB、XML、TXT等文件格式元数据文件及符合长期保存要求的文件格式全文文件的导入接收，给予唯一编号。导入时实现元数据、目录数据与对应电子文件、电子档案的自动关联。

数据导入过程中应支持数据的校验，如是否唯一、是否可以为空、日期格式是否正确等。当出现部分数据导入失败，应提供报告，指明哪些数据导入失败。如出现中断（如：断电、断网、死机等），应支持断点续传，再次导入时从中断记录处接续导入。

导入后形成符合《电子档案移交与接收办法》有关要求的交接凭据。

## 4. 电子文件和电子档案接收检测

能根据电子文件归档和电子档案接收的检测要求对接收的电子文件的“四性”和电子档案的有关属性进行检测，对检测不符合要求的电子文件或电子档案进行标记。检测完成后以件为单位形成检测业务实体元数据记录在元数据中。

# 二、整理功能

## 1. 组件

对来自不同系统的归档电子文件按照组件规则组成“件”，并对件内文档按一定的规则进行排序。

## 2. 分类

按照企业确定的分类方案，将归档文件进行分类，给定分类代号，并支持按年度、保管期限、机构（问题）等多种分类方案进行复合分类。

## 3. 划定保管期限

按照档案保管期限表对归档文件划定保管期限。

## 4. 组成保管单位

对归档文件按照一定规则自动或在人工干预下组成保管单位，并能按规则将保管单位内文件排序定位。能在人工干预下调整归档文件所属保管单位并重新排序定位。

## 5. 编号

依据档号编制规则形成档号，所形成档号唯一、简明、合理、稳定、可扩充。

## 6. 编目

提供规范的档案目录模板，能套打出案卷封面、案卷目录、卷内目录、归档文件目录、全引目录、备考表、卷（盒）封面和脊背等。应支持进行目录模板制作，用户可以对模板中的字体、打印内容、排序方式等进行调整。应提供打印预览功能，并支持多模板批量生成和打印功能。

根据模板对选定的归档电子文件或电子档案形成相应的目录。

支持案卷封面、卷内目录、归档文件目录或全引目录、案卷目录、备考表等通用格式文件导出功能。

## 7. 关联关系建立

具有建立同一件归档文件不同载体间或是同一归档文件不同子件间关联关系功能。

### 三、保管功能

#### 1. 档案整理

能依据电子档案保管和利用的业务要求分别建立相应数据分库。能根据全宗归属、分类号、保管单位序号及其整理规则、件号、件内文档排列规则等将电子档案排列定位和呈现。

#### 2. 电子档案长期保存

定期检测电子档案格式是否过时。根据所登记的格式列表，在启动检测时，给出格式过时预警提示。

对采用迁移、仿真、封装等方式保障数字档案信息的长期保管的技术使用过程形成相应的业务实体元数据。

#### 3. 电子档案存储格式转换与信息组织

支持采用插件技术实施电子档案存储格式的转换，将存入系统的电子档案转换为符合长期保存要求的存储格式。

系统可按《电子档案移交与接收办法》和离线存储载体容量进行信息组织，将组织好的档案保存至相应的离线存储载体上。

#### 4. 备份

应支持备份，包括软件系统备份、数据库备份和电子档案备份，能根据数据重要程度选择在线、离线等不同的备份方式。

#### 5. 定期鉴定

支持辅助人工完成档案的定期鉴定工作，根据保管期限、归档日期、密级等属性自动列出到期档案，提醒系统管理员或档案管理员等进行销毁前鉴定和保管期限调整、利用开放、密级调整等相关操作。

## 6. 销毁管理

支持对需销毁档案进行销毁申请、审批、建立销毁标记，实施监销、删除等操作。

## 7. 移交

具备根据进馆要求或其他有关要求生成移交电子档案信息包等功能，对已移交的电子档案或其他数字资源进行移交时间及去向标识。

## 8. 介质管理

支持对存放电子档案存储介质的统一管理，系统管理员可根据介质保管的实际需求为相应电子档案设置相应的存储介质。具有介质预警功能，当剩余容量达到设定阈值时，通知系统管理员。

# 四、统计功能

应支持按照全宗、分类、时间、文件格式、利用情况等设定规则进行统计、结果显示和打印，并以电子文件形式输出统计数据，支持自定义报表功能。统计结果能按《全国档案事业统计年报制度》给定格式输出统计数据。

# 五、利用功能

## 1. 检索

支持按属性检索档案，具备精确检索、模糊检索、高级组合检索、筛选检索、关联检索、深入二次检索等多种检索方式；有条件的企业可开发全文检索功能。

## 2. 浏览

应支持常见格式电子档案或其他数字资源的浏览，支持常见格式多媒体电

子档案的播放，应具有按权限下载功能。

### 3. 借阅

支持电子档案借阅申请、审批、授权，具有实体档案借阅预约、催还、归还等功能，审批流程符合档案利用管理制度。

### 4. 编研

支持用户根据实际需求设置编研专题，能自动将符合专题条件的档案进行归集。

### 5. 开放利用设置

支持对目录和全文设置开放利用标识。

### 6. 复制管理

具有电子档案复制申请、审批功能，能够进行复制；具有实体档案复制申请、审批、复制件分发登记等功能。

### 7. 利用效果反馈

支持用户对利用档案产生的经济效益或社会效益进行反馈。

### 8. 知识管理

可根据需要采用知识管理功能，支持档案信息的知识标签、知识地图、档案信息深层次加工、档案信息主动推送等。

## 六、系统管理功能

### 1. 日志管理

应有日志记载功能，记录系统启动和关闭信息、用户登录信息等，每条日

志至少应记录操作对象、用户、时间、计算机、操作类型等属性。提供针对日志的检索、审计、统计功能。

支持按照一定规则自动生成日志审计报告，支持按权限进行日志导出、删除及审批等操作。

## 2. 系统设置

支持对档案流程化管理，可根据不同类别档案管理要求，设置相应管理流程，可对流程进行跟踪与回溯。

支持代码表的设置维护，如年度、保管期限、密级、部门、分类号等代码；系统中的分类方案可根据用户需求增减。

支持元数据的设置维护，可以定义字段名称、类型、是否为空、是否唯一、字符串长度限制、取值范围、组合字段等信息。

具备全宗设置功能，能实现区分全宗进行管理。

## 3. 用户管理

支持直接录入用户信息和通过接口方式读取其他系统的用户信息，支持用户与 IP 地址绑定，支持“三员”分立。

## 4. 组织机构设置管理

支持直接录入组织机构信息，支持通过接口方式读取其他系统的组织机构信息，支持组织机构的多级管理功能。

## 5. 权限管理

具备权限管理功能，保证授权用户能够在其权限范围内进行合法操作。权限可以精细地控制指定用户对指定档案记录的访问和使用。系统应支持设置权限有效期，到期后系统自动取消权限。

## 6. 多全宗管理

可对纳入管理的每个立档单位建立全宗单位，支持全宗群管理，每个全宗单位有唯一全宗号、全宗名称。每个全宗单位的功能、档案类型、档案数据、

机构组织、用户、流程、权限等相对独立、互不影响。

## 7. 工作过程记录管理

应具备详细记录类似全宗卷信息的功能，以记录数字档案馆（室）管理过程。

# 七、传统载体档案辅助管理功能

## 1. 传统载体文件信息采集

具备传统载体文件目录著录和全文挂接功能，目录数据项在满足国家有关标准要求的前提下，可根据需要增减。目录数据采集时应具有校验功能。目录一旦著入能给予唯一的编号予以登记。

## 2. 传统载体文件归档整理

具备对目录信息与电子文件归档整理相同的整理功能。

## 3. 传统载体档案信息采集

具备传统载体档案目录著录和全文挂接功能，目录数据项在满足国家有关标准要求的前提下，可根据需要增减。目录数据采集时应具有校验功能。目录一旦著入能给予唯一的编号予以登记。

具备传统载体档案全文信息上传挂接功能，支持单个、批量文件上传等方式，能建立目录与其对应全文间的关联关系，并保持关联关系稳定。

## 4. 档案存放位置管理

能根据库房余量和档案的全宗归属、分类、档号等规则对传统载体档案提出存放位置建议，并在位置确定后进行排列定位。

## 5. 智能库房管理

有条件的单位可设置智能库房管理功能，根据现实库房虚拟出档案所在库

房、区域及档案柜架的列、节、层、排位等，实现对实体档案所在库房位置的管理；支持与智能密集架、PDA 设备、条码枪、电子标签等硬件设备的集成，实现实体档案的快速定位、出库、入库、盘点等操作；支持与温湿度监控设备的集成，实现档案库房的温湿度智能调控。

## 八、其他可选功能

有条件的企业可根据自身需要，增加诸如文件材料分发控制、工作计划进度管理、业务监督指导等功能。

## 第四节 模块设计

需求分析确定的功能是零散的，在实际开发中，需要按一定的逻辑进行组合，形成模块或分系统。因此，在软件开发中还需要进行模块设计。

### 一、模块或子系统划分的依据

一个合理的子系统或模块划分，应该是内部联系强，子系统或模块间尽可能独立，接口明确、简单，尽量适应用户的组织体系，有适当的共用性，即模块分解必须遵守的“耦合小，内聚大”的原则。按照结构化设计的思想，划分模块或子系统的依据有以下几种。

#### 1. 按逻辑划分

把类似的逻辑处理功能放在一个子系统或模块。例如，把“对所有档案数据采集”的功能放在一个子系统或模块，那么不管是文书档案、科技档案还是财务档案，数据录入都由这个子系统或模块实施。



## 2. 按时间划分

把同一时间段要执行的各种处理结合成一个子系统或模块。

## 3. 按过程划分

即按工作流程划分。从控制流程的角度来看，同一子系统或模块的许多功能都应该是相关的。

## 4. 按通信划分

把需要交互通信的处理结合成一个子系统或模块，这样可减少子系统间或模块间的通信，使接口简单。

## 5. 按职能划分

按职能划分子系统，再按逻辑划分模块。

## 6. 按用户类型划分

即按使用系统的用户类型来划分模块，如电子商务的购物平台，就按用户划分为购物用户使用的交易模块，供应商使用的供应商管理模块，平台管理用户使用的平台管理模块等。

# 二、模块划分实践

在电子档案管理系统中，模块或子系统的划分一般按照管理流程和用户类型来划分，其中按流程划分的比较多。

## 1. 按流程划分

按流程划分并组织功能形成模块，一般用于档案管理功能，见表 7-1。

表 7-1 档案管理信息系统模块划分表（按流程）

模 块	功 能
收集模块	电子文件手工登记
	电子文件在线接收登记
	电子文件离线接收
	传统载体文件信息采集
	传统载体档案信息采集
	电子文件和电子档案接收检测
整理功能	组件
	分类
	划定保管期限
	组成保管单位
	编号
	编目
	关联关系建立
	传统载体文件归档整理
保管功能	电子档案接收登记、检测
	档案整理
	电子档案长期保存
	电子档案存储格式转换与信息组织
	备份
	定期鉴定
	销毁管理
	移交
	介质管理
	传统载体档案存放位置管理
	智能库房管理
统计功能	
利用功能	检索
	浏览
	借阅
	编研
	开放利用设置

续表

模 块	功 能
利用功能	复制管理
	利用效果反馈
	知识管理
系统管理	日志管理
	系统设置
	用户管理
	组织机构设置管理
	权限管理
	多全宗管理
	工作过程记录管理

2. 按用户角色划分

按用户角色划分模块，见表 7-2。

表 7-2 电子档案管理系统模块划分表（按角色）

模 块	子 模 块	功 能
兼职档案员模块	收集模块	电子文件手工登记
		电子文件在线接收登记
		电子文件离线接收
		传统载体文件信息采集
		传统载体档案信息采集
		电子文件和电子档案接收检测
	整理功能	组件
		分类
		划定保管期限
		组成保管单位
		编号
		编目
		关联关系建立
		传统载体文件归档整理
档案管理员模块	保管功能	电子档案接收登记、检测
		档案整理
		电子档案长期保存

续表

模 块	子 模 块	功 能
档案管理员模块	保管功能	电子档案存储格式转换与信息组织
		备份
		定期鉴定
		销毁管理
		移交
		介质管理
		传统载体档案存放位置管理
		智能库房管理
	统计功能	
档案利用者模块	利用功能	检索
		浏览
		借阅
		编研
		开放利用设置
		复制管理
		利用效果反馈
		知识管理
系统管理员模块	系统管理	日志管理
		系统设置
		用户管理
		组织机构设置管理
		权限管理
		多全宗管理
		工作过程记录管理

## 第五节 界面及输入/输出设计

### 一、界面设计

用户界面是软件系统与操作者之间的接口，是控制和选择信息输入/输出的主要途径。用户界面设计需要通过规范化的定义分析用户需求，确定界面原型并能够让用户接受。

可用性是用户界面设计最重要的目标，主要从提高操作简单性、术语标准化与一致性、帮助功能和容错能力几方面进行考虑。

在进行用户界面设计时，需要对用户特性进行分析，了解软件所面向用户的经验和技能，要综合考虑用户的熟练程度，才能设计出适用于绝大多数用户所能接受的界面。概括起来，用户界面设计的主要工作就是规划界面空间布局、字体色彩协调、交互方式、输入/输出设计以及帮助与错误控制等。

界面设计应遵循如下原则。

#### （1）遵循 Windows 界面标准

包括合理窗体布局，选用标准控件，色彩搭配、字体设置，以及菜单、工具栏、状态栏、滚动条和右键菜单等。

#### （2）坚持一致性

确保一个系统不同窗体之间的控件使用具有一致性。

#### （3）提高可用性

标准的术语给用户以明确、必要的提示信息，不产生歧义或所指不明确；设计层次不同但相对完善的帮助提示；尽量以当前屏幕展示用户关心的信息，而不是新开窗口或让用户拖动滚动条来回游走，严重影响用户的工作效率；考虑用户环境差异，增强窗体、字体、图片显示对不同环境的适应性；增加鼠标悬停提示或状态栏提示；减少不必要的提示信息，尽可能让后台处理。

#### （4）追求完美与协调

界面布局大小适中，符合审美观点，能让人感觉舒适；将重要元素放于显著位置，遵循通常的阅读顺序布置屏幕元素；同一类型控件的合理分级与布局；进行人性化的色彩设计，坚持采用柔和、淡雅的色调，杜绝使用不合时宜的鲜艳颜色；恰当安排空白空间来突显界面元素，减轻视觉紧迫感；尽量选用标准 Windows 字体，坚持不同窗体中字体设置一致，字符间适当留下空白，消除空间局促性。

#### （5）完善容错与出错处理

增强容错机制，在输入大量数据的情况下，不因个别数据输入错误或其他故障而导致整批或整屏数据丢失，及时对输入的合法性进行检查，降低出错概率；对错误尽早地给出明确提示，指导用户迅速排除；尽量提供选择的方式代替手工输入，减少人工出错；尽量考虑到各种可能现象，不产生界面设计疏漏以致对系统构成危害。

## 二、输入/输出设计

### 1. 输入设计

输入设计的工作内容是：选择数据输入设备；输入数据格式的设计；输入数据正确性校验；联机系统的输入屏幕设计。

输入设计的原则包括以下五个方面。

（1）控制输入量。在输入时，只需输入基本信息，而其他可通过计算、统计和检索得到的信息则由系统自动产生。

（2）减少输入延迟。输入数据的速度往往成为提高信息系统运行效率的瓶颈，为减少延迟，可采用周转文件、批量输入等方式。

（3）减少输入错误。应采用多种输入校验方法和有效性验证技术，减少输入错误。

（4）避免额外步骤。

（5）输入过程应尽量减化。输入过程要简单易用，不能因为查错、纠错而使输入复杂化，增加用户负担。

### 2. 输出设计

输出应考虑输出方式，输出设备，输出介质，输出报表信息和图形及多媒体信息等。

## 第八章 电子档案管理系统实施

### 第一节 电子档案管理系统实施概述

所谓“系统实施”，是指将系统设计阶段的结果在计算机上实现，将原来纸面上的、类似于设计图纸的系统方案转换成可执行的应用软件系统。在此过程中，需投入较多人力、物力，并花较长时间进行硬件购置与安装、软件采购、系统调试、人员培训、系统转换、系统管理等一系列工作，在系统生命周期中将这个阶段称之为系统实施阶段。

#### 一、系统实施目标

在系统分析与设计阶段，开发人员设计了电子档案管理系统的逻辑模型和物理模型。系统实施阶段的目标就是把系统的物理模型转换成可实际运行的信息系统。

#### 二、系统实施技术原则

系统实施的基本技术原则为：最大限度地利用现有硬件及网络资源，适当地对业务流程进行重组，以达到优化管理的目的。

#### 三、系统实施的主要内容与步骤

系统实施是一项复杂性工程，信息系统的规模越大，实施阶段的任务就越复杂。电子档案管理系统涉及企业多个部门和人员，并以档案、文件以及其他数字资源的日常管理为对象，实施任务复杂而繁重。一般来说，系统实施阶段

工作主要有以下几个方面。

- (1) 按总体设计方案部署计算机网络系统。
- (2) 建立数据库系统。
- (3) 软件采购。
- (4) 导入基础数据，进行系统测试。
- (5) 人员培训，系统转换，试运行。
- (6) 系统的评价与验收。

其中，软件采购作为大多数企业获得电子档案管理系统软件的主要方式，本章将重点阐述，自主开发方式参见第五章。

## 第二节 软件采购

软件采购看似简单，但实际上许多单位采购的软件根本无法使用，或达不到预期效果，其主要原因是没有做好采购前期的工作。软件采购的主要工作有：初始需求分析、原型测试、商务谈判、签订合同及付款、客户化开发、系统安装等。

### 一、初始需求分析

软件采购时的初始需求分析，是确定待选软件厂商范围的需要，只有经过初始需求分析，才能确定适合本企业系统软件的厂商范围，初始需求分析的内容如下。

- (1) 明确软件采购的目的是升级还是全新版本。如果是升级，则只能选择原软件厂商；如果是全新版本，则选择范围较广。
- (2) 提出软件技术需求，包括开发平台需求、技术架构需求等。
- (3) 提出软件功能需求，具体参见第5章。
- (4) 提出软件性能需求，具体参见第5章。
- (5) 收集可供选择的软件厂商，并对其产品特点进行分析。
- (6) 初步确定待选厂商范围。



## 二、原型测试

确定软件初始需求后，应对纳入待选厂商范围及厂商产品原型进行逐一测试，严格对照初始需求，找出原型软件在功能、性能与初始需求的差距，并逐一记录，形成原型测试报告。

经过原型测试，企业可对各厂商的产品是否符合初始需求作出初步判断，这对后续商务谈判有重要参考价值。测试过程中，还可集中各厂商产品优点，并纳入初始需求中。

## 三、商务谈判

企业通过原型软件测试，对各厂商产品有了初步了解，接下来将选定一家厂商的产品作为采购对象。软件采购应采取招投标的方式，并严格按照国家和企业有关招标采购的规定执行。其中，要特别注意招标书的编制。为避免后续争议，招标书要对需求尽量详尽明确，需求表述不使用有歧义性语句。当然，软件购置除考虑产品功能外，软件厂商的信誉、经验、资质等也是要考虑的重要因素。

## 四、签订合同及付款

经招投标程序确定厂商后，采购进入签订合同和付款阶段。软件采购合同一般分为商务和技术协议两个部分，技术协议可作为合同附件成为合同的一部分，与合同具有同等法律效力。技术协议一般约定软件功能需求、性能需求等内容，也可包括软件的技术架构、开发平台约定等。

## 五、客户化开发

一般来说，软件厂商已开发的标准版或演示版系统与企业需要购买的目标系统存在一定差距，特别是功能差距。由于软件厂商演示的系统是为其他用户所开发，不一定完全满足或符合本企业目标系统的需求，因此，还需对系统

进行二次开发，或称客户化开发。

客户化开发与新系统开发类似，只是客户化开发可以在标准版系统的基础上，通过需求补充分析、设计、实施等过程实现。在这一过程中，需求分析是重点，用户要大量参与，并与系统开发人员一起分析标准版系统与目标系统的差距，提出二次开发需求，并采用生命周期法不断更新，直至达到用户满意，主要过程如图 8-1 所示。

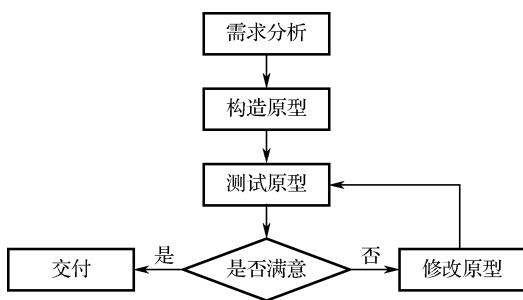


图 8-1 客户化开发流程图

## 六、软件安装

经客户化开发并测试后，满足合同约定功能要求的软件，可在企业服务器上安装。与前述工作相比，安装过程较为简单。但个别软件需安装在特殊环境下，企业信息部门要了解特殊环境需求，并提供相应环境。由于软件安装工作较少反复，软件安装的同时可组织本企业运维人员进行软件安装培训，这也为企业后续的软件运维工作奠定基础。

安装软件后，进入基础数据准备阶段，为软件试运行做好准备。

### 第三节 基础数据准备

#### 一、用户数据准备

用户数据是电子档案管理系统的重要基础数据，主要包括用户角色及权限

设置、用户编码设置、用户数据建立等。

## 1. 用户角色及权限设置

### (1) 系统管理用户设置

### (2) 用户角色设置

① 按流程设置角色，即按照档案数据管理流程设置不同的角色，如数据录入角色、数据审核角色、数据过程管理角色、数据检索角色等。

② 按权限设置角色，即按照数据的操作权限设置不同的角色，如档案管理员、兼职档案员、档案检索人员

③ 按功能设置角色，即将档案管理信息系统用户分为对流程施加影响角色和对信息施加角色。

#### ● 对流程施加影响角色。

可设置流程发起角色、流程审核角色、流程批准角色和流程办理角色。如档案管理信息系统有借阅管理流程，是借阅人提出申请，经部门领导审核和档案部门批准，最后由档案管理员实施。

#### ● 对信息施加影响角色。

指对具体档案信息数据记录施加影响的程度来设置角色，可设置数据录入角色、数据修改角色、数据利用角色，它们的权限区别是：数据录入角色是录入数据，但无数据修改和删除权，数据修改权主要是针对数据审核与档案整编人员，它只能对数据的个别字段进行修改，而数据利用角色则无权对数据进行任何修改。

④ 按菜单功能设置角色，即用户不再区分角色，而是根据权限予相应的菜单操作权限。

### ⑤ 按可操作数据或字段范围设置角色。

可设置以下角色：

- 文书档案管理员：只能对文书档案数据进行操作；
- 产品档案管理员：只能对产品档案数据进行操作；
- 基建档案管理员：只能对基建档案数据进行操作；
- 设备档案管理员：只能对设备档案数据进行操作。

或设置以下角色：

- 归档文件录入角色：对基本字段具有操作权限；
- 档案鉴定员角色：对保管期限字段有操作权限；

- 档案整理员角色：对档号、分类号、卷内序号等字段具有操作权限；  
档案检索者角色：不对任何字段具有操作权限。
- 上述系统角色设置可单一使用，也可综合运用，将角色设置更加合理化。

## 2. 用户编码设置

用户编码可单独编码，可采用本企业统一编码体系。

## 3. 用户数据建立

通过批量倒入的方式建立用户数据。

# 二、数据字典建立

数据字典是系统运行的基础，必须在系统测试之前建立。具体包括档案分类代码、部门机构代码、密级代码、保管期限代码等。

## 1. 档案分类代码

档案分类代码主要是为系统中的电子档案标明类别代号，它须与企业档案分类法、分类方案或分类表相一致。

## 2. 部门机构代码

部门机构代码可按照本企业部门机构代码表执行，如 BGT 办公厅。

## 3. 密级代码

密级代码按照密级从高到低的顺序，一般设置如下。

- 001 绝密级：同国家绝密级（使用较少，绝密级档案不允许数字化）；
- 002 机密级：同国家机密级；
- 003 秘密级：同国家秘密级；
- 004 内部秘密级：企业内一定范围人员有权查阅；
- 005 内部一般级：企业所有人员有权查阅，但不得在媒介公开；
- 006 内部公开级：可在企业内部媒介公开；

007 外部级：可在企业外部媒介公开。

#### 4. 保管期限代码

设置保管期限代码比较简单，按国家有关规定，可将档案保管期限设置为永久、30 年、10 年。

由于有些档案采用三分法，还需设置为长期、短期。

有的档案保管期限采用不是 30 年、10 年期的标时制，如设备报废后 5 年，产品报废后 5 年等代码。

### 三、界面设计

目前大多数电子档案管理系统都具备界面自定义功能，在系统实施时，需对界面进行自定义，以满足用户的数据输入需求。界面定义主要是针对数据著录界面，其他界面一般不涉及。档案数据著录是使用最频繁、最影响档案信息系统工作效率的界面，必须对它进行认真定义。定义档案数据著录界面应从以下几方面考虑。

- (1) 界面要美观大方，输入框布局合理，选择显示的字体、字号适中。
- (2) 标签定义准确，尽量使用档案专用术语。
- (3) 输入项先后顺序的安排，要充分考虑实际来源，而不是以著录规范中的先后次序为序。
- (4) 代码编码排序，要按使用频次排序，减少反复使用鼠标点击的频次。

### 四、报表设计

档案工作中有大量的报表需要使用，报表设计也是系统实施中的重要工作内容，需要设计的档案工作报表主要如下。

- (1) 卷内目录（按卷管理）和归档文件目录（按件管理）。
- (2) 案卷内录。
- (3) 档案移交清单。
- (4) 档案统计年报。

- (5) 备考表。
- (6) 案卷封面与脊背标签。
- (7) 自定义报表。

## 五、分类结点树方案设计

数字档案馆中电子档案的组织方式，是按照一定分类逻辑进行组织的，因此落实到软件界面中，一般在软件左侧建立分类结构树，这个分类结构树是将电子档案管理系统中档案信息进行组织的关键。为建立逻辑合理的分类结点树，事先有必要设计方案，建立遵循的分类规则，否则档案信息大量交叉会影响电子档案管理系统操作人员的理解，影响管理效率。一般是用进行了必要调整的本企业档案分类方案作为分类结点树建设方案，其中不需确定结点树的层次。图 8-2 是某企业设计的结点树方案。

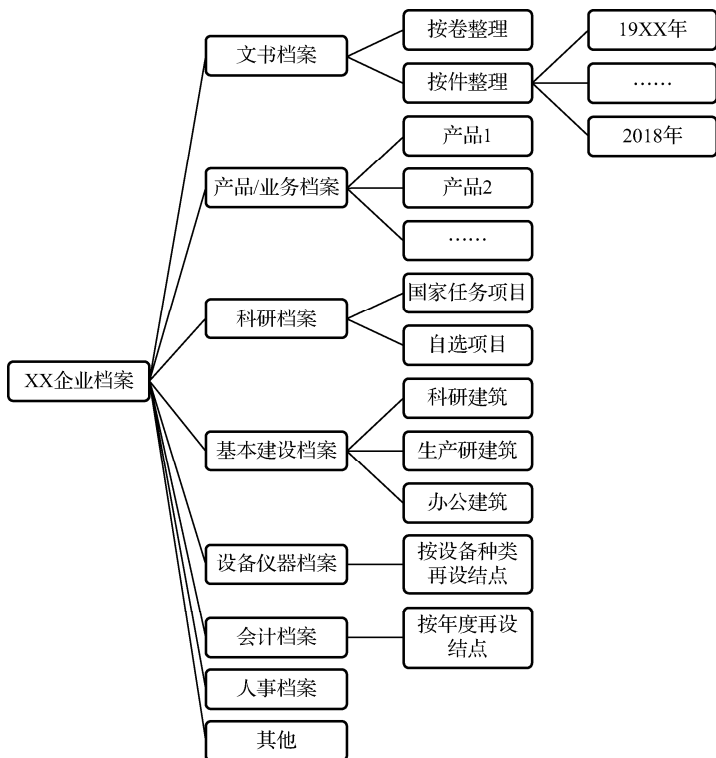


图 8-2 某企业档案结点树图

## 第四节 系统测试

### 一、系统测试概述

#### 1. 测试目的

对系统进行测试的目的，是尽可能多触发系统中的错误和缺陷。必须纠正的一个错误看法是：认为测试是为了证明程序的正确性，即测试是为了证明程序不再有错误，事实证明这是不现实的。因为要通过测试来发现程序中的所有错误需要穷举所有可能的输入数据，检查它们是否产生正确结果。例如，一个需要三个 16 位字节长的整型输入程序，输入数据的所有组合情况大约有 30 多万亿种，若每组数据的测试时间为 1ms，大约需要 1 万年时间才能完成。

#### 2. 测试用例

进行系统测试，除了准备测试数据外，还应确定该组数据测试正确时的输出结果，可称它为预期结果。测试时，将实际的输出结果与预期结果相比较，若相同则表示程序正确，反之则表示系统出现了错误，因此测试用例是由测试数据和预期结果构成的。

为了发现程序中的错误，应尽力设计能暴露错误的测试用例。一个好的测试用例是极有可能发现尚未发现的错误的。一次成功的测试是发现了迄今为止尚未发现的错误的测试。

#### 3. 测试原则

(1) 设计测试用例时，必须确定预期输出结果。如果一个测试用例只有测试数据而无预期结果，则无法判断测试结果是否正确。

(2) 交叉测试原则。系统测试时应避免由开发人员测试自己编写的程序，这是因为程序错误往往是由于开发人员对问题说明的误解造成的，自己测试本人程序就不易找出因这种误解而产生的错误。此外，系统开发是一项建设性工

作，而测试是一项破坏性工作，这对于开发人员来说，在心理上是难以容忍的。为了证明程序的正确性，开发人员往往不去选择容易发生错误的测试用例，而选择容易通过的测试用例。一般来说，开发人员只在模块测试时可以测试本人编写的程序。

(3) 彻底检查每一个测试结果，防止遗漏掉错误。

(4) 对非法的和非预期的输入数据也要与合法的和预期的输入数据一样编写测试用例。例如，判断输入的三个数据是否构成等腰三角形、等边三角形以及不等边三角形的程序。测试这个程序时，不仅要选“5, 6, 7”和“7, 7, 7”这样一些合理数据作为测试用例，而且还要选用“1, 2, 3”和“1, 2, 4”这样一些不合理的输入数据，以证实程序不会把这些不构成三角形的边长错误地认为是“不等边三角形”。

(5) 在测试中除了要检查程序正确与否外，还应检查程序是否有效。

(6) 测试用例存档。测试用例除了具有测试作用外，在系统运维中的改正错误或程序维护后，还要进行重新测试，因此，测试用例应存档以备后续运维时重用。

## 二、测试方法

程序测试方法主要有白盒测试法和黑盒测试法。

### 1. 白盒测试法

白盒测试是把程序看成装在一只有透明的白盒子里，测试者完全了解程序的结构和处理过程。它根据程序的内部逻辑来设计测试用例，检查程序中的逻辑通路是否都按预定的要求正确地工作。

### 2. 黑盒测试法

黑盒测试法是把程序看成一只黑盒子，测试者完全不了解（或不考虑）程序的结构和处理过程。它根据规格说明书规定的功能来设计测试用例，检查程序的功能是否符合规格说明书的要求。



### 三、测试类型

系统测试可分为四种类型，即模块测试、联合测试、验收测试和系统测试。

#### 1. 模块测试

模块测试是对一个模块进行测试，根据模块的功能说明，检验模块是否有错误。这种测试在各模块编程后进行。模块测试通常采用白盒测试。

模块测试一般由编程人员自己进行，具体测试内容如下。

- (1) 模块界面：调用参数数目、顺序和类型。
- (2) 内部数据结构：如初始值的正确性、变量名称的一致性、共用数据的正确性等。
- (3) 独立路径：是否存在不正确的计算、循环及判断控制。
- (4) 错误处理：预测错误的产生及后处理，看是否和运行一致。
- (5) 边界条件：对数据大小界限和判断条件的边界进行跟踪运行。

#### 2. 联合测试

联合测试即通常说的“联调”，它是对由各模块组装而成的程序进行测试，主要检查模块间的接口和通信，可以发现总体设计中的错误。因为各个模块单独执行可能无误，但组合起来可能相互产生影响，而会出现意想不到的错误，所以，应采用联合测试发现模块组合后可能出现的错误。联合测试方法一般有两种，即根据模块结构图由上到下或由下到上进行测试，联合测试通常采用黑盒测试。

#### 3. 验收测试

验收测试是检验系统说明书和各项功能与性能是否与用户的需求一致，它是以需求说明书为依据测试的。验收测试通常采用黑盒测试。

验收测试的方法，一般列出一张清单，左边是需求的功能，右边是发现的错误或缺陷。常用的验收测试有 $\alpha$ 测试和 $\beta$ 测试。 $\alpha$ 测试是在开发人员现场由客户来实施的，被测试的系统是开发人员从用户的角度进行常规设置的环境下运

行的； $\beta$ 测试是在一个或多个客户的现场，由软件的最终用户实施的。与 $\alpha$ 测试不同的是， $\beta$ 测试时开发人员通常不在场。 $\alpha$ 测试和 $\beta$ 测试除了进一步发现程序中的错误外，还能发现使用上的问题。

#### 4. 系统测试

系统测试是对整个系统的测试，将硬件、软件和操作人员看作一个整体，检验它是否有不符合系统说明书的地方。这种测试可以发现系统分析和设计中的错误。例如，安全测试是测试安全措施是否完善，能否保证系统不受非法侵入。

### 四、黑盒测试用例设计

黑盒测试是根据规格说明所规定的功能来设计测试用例，它不考虑程序内部结构和处理过程。常用的黑盒测试技术有等价类划分、边界值分析、错误猜测和因果图等。

#### 1. 等价类划分

在进行测试时，不能穷举所有可能的输入数据来进行测试，所以只能选取少量有代表性的输入数据来查找尽可能多的程序错误。

##### (1) 确定等价类

根据程序的功能说明，对每一个输入条件（通常是说明中的一句话或一个短语）确定若干有效等价类和若干无效等价类，可以使用表格形式分别写出输入条件、有效等价类和无效等价类。在确定等价类时，可考虑下列规则。

① 如果某个条件规定了值的范围，那么可确定一个有效等价类（输入值在此范围内）和两个无效等价类（输入值小于最小值或大于最大值）。例如，规定考试成绩在 0~100 之间，则有效等价类为： $0 \leq \text{成绩} \leq 100$ ，无效等价类为：成绩 $<0$  或成绩 $>100$ 。

② 如果某个输入条件规定了值的个数，那么可确定一个有效等价类和两个无效等价类。例如，规定每个运动员的参赛项目数为 1~3 项，则有效等价类为  $1 \leq \text{成绩} \leq 3$ ，无效等价类为项目数 $<0$  和项目数 $>3$ 。

③ 如果某个输入条件规定了一个输入值的集合（即离散值），而且程序对

不同的输入值做不同的处理，那么每个允许的值定为一个有效等价类，另外还有一个无效等价类（任意一个不允许的输入值）。例如，规定考试成绩为优、良、中、及格和不及格，则可确定 5 个有效等价类和无效等价类。

④ 如果某个输入条件规定了输入数据必须遵循的规则（如标识符的第一个字符必须是字母），那么可以确定一个有效等价类（符合此规则）和若干个无效等价类（从各种不同的角度违反该规则）。

⑤ 如果某个输入条件规定输入数据是整形，那么可以确定 3 个有效等价类（正整数、零、负整数）和一个无效等价类（非整数）。

⑥ 如果某个输入条件规定处理的对象是表格，那么可确定一个有效等价类（表有一项或多项）和一个无效等价类（空表）。

以上列举的只是一些规则，实际情况往往千变万化，在遇到具体问题时，可参照上述规则的思想来划分等价类。

## （2）设计测试用例

利用等价类设计测试用例的步骤如下。

① 为每个有效等价类和无效等价类编号。

② 设计一个新的测试用例，以尽可能多地覆盖尚未被覆盖的有效等价类。重复这一步直至所有的有效等价类都被覆盖为止。

③ 为每个无效等价类设计一个新的测试用例。

在测试时，当一个测试用例发现了一个错误时，往往就不再检查这个测试用例是否还可能发现其他错误。无效等价类都是测试非正常输入数据的情况，因此每个无效等价类都很有可能查出程序中的错误。所以，要为每个无效等价类设计一个新的测试用例。

## 2. 边界值分析

经验证明，程序往往在处理边缘情况时易出现错误，因此检查边缘情况的测试用例效率是比较高的。边界条件是指相对于输入与输出等价类直接在其边界上，或稍高于其边界，或稍低于其边界的这些状态。例如，某个输入条件说明值的范围是 0~100，则可以选 0、100、-1、101 为测试用例。再如，一个输入文件可以有 1~128 个记录，则分别设计有 0 个、1 个、128 个、129 个记录的输入文件，等等。使用等价类划分方法设计测试用例时，原则上来讲，等价类中的任一输入数据都可以作为该等价类的代表用作测试用例。而边界值分析

则是专门挑选那些位于边界附近的值作为测试用例。

下面是一些用边界值分析方法设计的测试用例。

(1) 新记录

- ① 在文件第一个记录之前加一个记录。
- ② 在文件最后一个记录之后加一个记录。
- ③ 插入的新记录对应实体是实际不可能存在的。
- ④ 记录的域不全。

(2) 处理业务

- ① 处理文件的第一个记录。
- ② 处理文件的最后一个记录。
- ③ 处理中间的一个记录。
- ④ 处理同一程序刚建立的记录。
- ⑤ 连续处理相邻的记录。
- ⑥ 试图处理一个不存在的记录。
- ⑦ 处理业务使某个数值超过常规（如学习成绩为负）。
- ⑧ 对某些关键数据输入有错误的数据。
- ⑨ 同一业务处理过程中造成多重例外和出错。

(3) 记录删除

- ① 删除文件的第一个记录。
- ② 删除文件的最后一个记录。
- ③ 试图删除不存在的记录。
- ④ 连续删去多个记录。
- ⑤ 删去一个记录，并试图处理这个记录。

(4) 试验逻辑

- ① 检查所有能产生最大值、最小值和平均值的计算。
- ② 除式中除数为零。
- ③ 数据域放最小数或最大数。
- ④ 数据域填入允许值之外的数。

(5) 报告程序

- ① 负号是否全打印出来。
- ② 全 9 是否打全。
- ③ 全 0 时查看高位压缩情况。

- ④ 交叉结算平衡结果是否有报告。

### 3. 错误猜测

错误猜测是一种凭直觉和经验推测某些可能存在的错误，从而针对这些可能存在的错误设计测试用例的方法。这种方法没有机械的执行步骤，主要依靠直觉和经验。

例如，测试一个排序子程序，可考虑以下情况。

- (1) 输入表为空。
- (2) 输入表只有一个元素。
- (3) 输入表的所有元素都相同。
- (4) 输入表已排好序。

### 4. 因果图

边界值分析和等价等方法都没有考虑输入条件的各种组合，在输入条件的组合情况数目相当大的情况下，应该用某种方法来选择输入条件的子集，再考虑它们的组合。因果图是一种帮助人们系统地选择一组高效测试用例的方法，其特点如下。

- (1) 考虑输入条件之间的组合关系。
- (2) 考虑输出条件对输入条件的依赖关系，即因果关系。
- (3) 测试用例发现错误的效率高。
- (4) 能检查出功能说明中的某些不一致或遗漏。

采用因果图设计测试用例的步骤如下。

(1) 分割功能说明书。对于规模较大的程序来说，由于输入条件的组合数太大，所以难以使用因果图。可以把它划分成若干部分，然后分别对每个部分使用因果图。例如，测试某语言的编译程序时，可以把该语言的每个语句作为一个部分。

(2) 识别出“原因”和“结果”，并加以编号。“原因”是指输入条件的等价类，“结果”是指输出条件或系统变换。例如，更新主文件就是一种系统变换。每个原因或结果都对应于因果图中的一个节点。当原因或结果成立（或出现）时，相应的节点取值为 1，否则为 0。

(3) 根据功能说明中规定的原因和结果之间的关系画出因果图。

(4) 根据功能说明在因果图中加上约束条件。其约束条件有互斥、包含、唯一、要求、屏蔽。

(5) 根据因果图画出判断表。

(6) 为判断表的每列设计一个测试用例。

## 五、排错

测试是为了发现程序中存在错误,排错是确定错误位置和性质并改正错误。排错的关键是找到错误的具体位置,一旦找到,修正错误就相对容易。确定错误的方法有以下几种。

### 1. 试探法

分析错误的外在表现形式,猜想程序故障的大概位置,采用一些简单的纠错技术,获得可疑区域的有关信息,判断猜想是否正确,经过多次试探找到错误的根源。这种方法与个人经验有很大关系。

### 2. 跟踪法

对小型程序,可采用跟踪法,跟踪法有正向跟踪法和反向跟踪法。正向跟踪是沿着程序的控制流,从头开始跟踪,逐步检查中间的结果,找到最先出错的地方;反向跟踪是从发现症状的地方开始回溯,沿着控制流往回人工追踪程序代码,直到确定错误根源。

### 3. 对分查找法

若已知程序中的变量在中间某点的正确预期值,则可以用赋值语句把变量置成正确值,运行程序查看输出结果是否正确,若输出结果没有问题,则说明程序错误在前半部分,否则在后半部分。然后对有错误的部分再用这种方法,逐步缩小查错范围。

### 4. 归纳法

从错误出现的线索出发,分析这些线索之间的关系,确定错误的位置。首

先要收集、整理程序运行的有关数据，分析出错误规律，在此基础上提出错误的假设，若假设能解释原始测试结果，说明假设得到证实；否则重新分析，提出新的假设，直到最终发现错误原因。

### 5. 演绎法

分析已有的测试结果，设想所有可能的错误原因，排除不可能的、互相矛盾的原因。对剩下的原因，按可能性大小，逐个作为假设解释测试结果，直至找到错误原因。必要时，对列出的原因加以补充修正。

## 六、采购测试

电子档案管理系统的开发测试一般由系统开发商完成，但档案部门在购买电子档案管理系统时仍要进行大量的测试工作，以验证系统符合要求。电子档案管理系统采购时的测试称之为采购测试，采购测试主要有以下工作。

### 1. 功能测试

首先对各子系统进行独立测试，验证是否完成了基本功能，子系统内逻辑上是否存在错误；然后再验证各子系统的健壮性和容错能力；最后再以仿真的形式，按照基本事件、备选事件流，验证流程的正确性。对子系统进行独立测试时，需要按照事件流的顺序进行测试，以方便测试数据的传递。

各测试人员独立工作并共享相同的数据或数据库，因此要求测试人员的测试数据必须严格按照测试用例中准备的数据输入，不得随便输入或修改其他测试工程师所使用的数据，否则，可能会造成其他测试工程师的测试结果无效。

最后，测试软件和硬件接口间是否存在问题，主要测试数据导入、导出，打印功能是否正常。

### 2. 性能测试

性能测试需要模拟多用户同时访问服务器时，应用系统的压力及响应时间等。

#### (1) 测试场景设计

电子档案管理系统共有多个应用模块，由于不同的系统模块访问量不同，

因此需要设计几个并发访问量相对较大的模块和功能，模拟并发使用的场景，进行性能压力测试。

场景设计思想是：逐步提高系统用户同时并发登录，并发下载数据，以检查系统的长期稳定性。

## （2）性能测试方法

### ① 单业务测试。

单一业务测试主要针对单一业务流程设计，主要考察某一项单一业务在各种情况下的响应时间、系统资源占用、事务成功率等指标。对设定的每个单一场景，将分别进行 10 用户、20 用户、30 用户、40 用户、50 用户的同时并发模拟测试，同时考虑加载所有并发用户，实行多次循环。

### ② 混合业务测试。

混合业务场景模拟系统真实环境而设计，主要为了测试整个系统在各种情况下的响应时间、系统资源占用、事务成功率等指标。测试时采用大用户量并发方式混合测试真实情景下的效果。加载并发用户时，要考虑系统真实投入使用时可能的并发用户数。

### ③ 疲劳测试。

疲劳测试是采用系统稳定运行情况下能够支持的最大并发用户数，持续执行一段时间业务，通过综合分析交易执行指标和资源监控指标来确定系统处理最大工作量强度性能的过程，同时加载所有并发用户，持续运行 8 小时。

## （3）测试内容

- ① 一定数据量，一定并发用户数，系统登录响应时间。
- ② 一定数据量，一定并发用户数，档案查询响应时间。
- ③ 一定数据量，一定并发用户数，档案管理业务操作响应时间。
- ④ 一定数据量，一定并发用户数，档案全文上传响应时间。
- ⑤ 一定数据量，一定并发用户数，档案全文下载响应时间。
- ⑥ 一定数据量，一定并发用户数，系统档案数据汇总响应时间。

## 3. 安全保密测试

安全保密性测试主要测试系统在安全保密方面的功能和性能，主要测试要点如下。

### （1）用户权限管理功能。



(2) 身份验证功能。系统提供用户身份验证，身份验证不成功有次数限制及处理措施。

(3) 密码设置功能。系统提供密码设置策略，包括有效期、最小长度、复杂度、非空设置、大小写敏感度等；系统所有密码不得明码存储和传输。

(4) 数据传输安全。系统对于有特殊安全要求的数据，应在传输过程中进行必要的加密处理，如 UIL 参数信息、涉密文件信息、授权信息、个人密码信息等。

(5) 容错性。系统可以屏蔽用户误操作，系统对错误有正确提示，进行错误操作时，系统不崩溃、不异常退出、不丢失数据。

(6) 数据校验。系统应对数据的有效性、数据项之间的逻辑关系进行校验，保证数据有效；应保证数据的完整性和一致性，不会因删除或反复的更新而被破坏或留下垃圾数据；对不符合要求的输入数据，系统应使用中文给出简洁、准确的提示信息，必要时应给出帮助。

(7) 数据备份与还原。系统自身提供数据备份与还原手段，系统应设置备份周期及备份策略。

(8) 日志和审计。系统对关键数据的变更应记录日志，系统提供日志信息分类、查询、统计、分析、导出等管理功能，提供安全审计功能。

(9) 超时自动退出。超过一定时限未进行操作，系统可以自动退出。

#### 4. 兼容性测试

兼容性测试主要是测试电子档案管理系统在不同软硬件环境中运行的可能性，购买电子档案管理系统时必须对系统可能运行的不同软硬件环境进行测试，以验证其兼容性。

#### 5. 用户文档测试

测试时可有意根据用户文档进行，以检验用户文档是否完整、准确。

### 第五节 人员培训

为了使电子档案管理系统能够按预期目标正常运行，对用户人员进行必要

的培训是一项不可忽视的工作。

电子档案管理系统是一个人机系统，它的正常运行需要很多人参与工作，并承担系统所需输入信息的人工处理和计算机操作过程。这些人员通常来自业务管理部门和职能部门，他们熟悉或精通原来的人工处理过程，但缺乏计算机处理的有关知识，为了保证系统顺利使用，必须提前培训有关人员，需要培训的人员主要有以下三类。

## 一、业务和职能部门人员

电子档案管理系统能否顺利运行，并达到预期目标，在很大程度上与企业各业务部门和职能部门的人员有关。因此，可以通过讲座、报告会的形式，向他们说明电子档案管理系统的目标、特点和功能，说明系统的结构及运行过程，以及对企业组织机构、工作方式等产生的影响。对这些人员进行培训时，必须做到通俗、易懂、具体，尽量不采用与实际业务领域无关的计算机专业术语，培训的内容主要有以下几方面。

- (1) 电子档案管理系统的功能。
- (2) 电子档案管理系统对业务人员的技术要求。
- (3) 电子档案管理系统的应用，给现有工作流程带来的改革与变化。
- (4) 电子档案管理系统实施后，工作任务的效率评价。

## 二、档案管理人员

档案管理员是电子档案管理系统的直接使用者，统计资料表明，电子档案管理系统在运行期间发生的故障，大多数是由于使用方法错误造成的。所以，档案管理员的培训应该是人员培训工作的重点。

应该提供比较充分的时间对档案管理员进行培训，除了学习必要的计算机专业知识外，还须向他们传授新系统的工作原理、使用方法、简单出错的处置等知识。一般来说，在系统调试阶段就可以让档案管理员一起参加培训。例如，在输入基础数据、程序调试操作过程中，对他们熟悉电子档案管理系统，极有益处。

### 三、系统维护人员

对系统维护人员来说，要求具有一定的计算机专业知识，并对电子档案管理系统的原理和维护知识有深刻理解，系统维护人员一般由计算机中心和档案部门的计算机专业人员担任。

有条件时，应该请系统维护人员和档案管理员，或其他与电子档案管理系统有直接接触的人员，如机要人员、兼职档案管理员、基层资料管理员等，与软件供应商的实施人员进行充分交流与讨论，有助于他们了解整个系统的全貌，并为日后的系统应用打好基础。

对于大中型企业或部门用户，人员培训应列入企业或部门培训计划，并在执行时进行合理的分工与协作，培训部门负责培训的组织与协调，具体实施人员负责教材资料的编制、讲稿准备及考核标准的制定。

## 第六节 系统切换

在电子档案管理系统实施工作中，完成了系统安装及人员培训后，接下来就是系统的启动和运行。如果是全新的系统，在系统投入运行前期必须在手工管理（包括旧系统）与电子档案管理系统间进行平滑的过渡，也即系统切换。为了保证原手工管理的工作有条不紊地、顺利地转移到新系统，在系统转移前应仔细拟订方案和措施，确定具体的步骤。

系统的转换方式通常有三种：直接转换、平行转换、分段平行转换。

### 一、直接转换

直接转换就是选定某一时刻，停止档案管理业务的手工管理或旧系统运行，将新的电子档案管理系统立即投入使用，全部档案业务工作转移到新系统中进行，中间没有过渡阶段，见图 8-3。采用这种方式可节省一定人力物力，适用于新系统功能应用不太复杂，或旧系统完全不能使用的场合。但是，这种转换

方式风险较大。系统转换时还应做好准备，一旦新系统达不到预期目的，可立即采取相应措施予以补救。



图 8-3 系统直接切换方式图

## 二、平行转换

平行转换就是手工管理或旧系统与新的电子档案管理系统平行运行一段时间，即在手工管理或旧系统运行的同时试运行新系统。经过试运行后，如已达到预期目标，再用新系统替换手工管理或旧系统。这种转换方式在平行运行期间，使手工管理或旧系统运行与新系统运行并存，一旦新系统出现问题，即可暂行停止而不会影响正常工作，见图 8-4。



图 8-4 系统平行切换图

平行转换首先以原有手工管理或旧系统运行为正式作业，新系统的处理结果作为校核使用，直至最后终止手工管理或旧系统运行。根据系统的复杂程度、规模大小和基础条件建设程度等不同，平行运行的时间一般在三个月到一年之间。

采用平行转换的风险较小，在转换期间还可比较原有手工管理或旧系统运行与新系统运行的性能、效率等，并让有关人员得到全面培训。因此，平行转换是一种比较通用的方式。

但在平行运行期间，由于两种处理方式并存，使成本较高，并对工作效率有一定影响，这就需要做好周密的计划并加强管理。

三、分段平行转换

分段平行转换，是对手工管理或旧系统与新系统在业务功能上分期分批进行转换的一种方式，即手工管理或旧系统先与新系统的第一个子系统平行运行一段时间，如达到预期效果，则用新系统第一个子系统替换手工管理或旧系统中可由新系统第一个子系统执行的功能，同时试运行新系统的第二个子系统，再使手工管理部分业务或旧系统部分功能与新系统第二个子系统平行运行一段时间，如也达到预期效果，则用新系统第二个子系统再替换手工管理或旧系统中可由新系统第二个子系统执行的功能，同时试运行新系统的第三个子系统，以此类推，直至新系统全部替代手工管理或旧系统。这种方式集中了上述两种方式的优点，它既降低了系统切换风险，保证了系统平稳运行，成本也较低，具体见图 8-5。

手工管理或旧系统运行		手工管理部分业务或旧系统部分功能运行	
	新系统子系统1 试运行	新系统子系统1 运行	
		新系统子系统2 试运行	新系统子系统2 运行
.....			
			新系统子系统n 试运行
			新系统子系统n 运行

图 8-5 系统分段平行切换示意图

采用分段转换时，手工管理或旧系统与新的电子档案管理系统的转换次序及具体步骤，应根据实际情况灵活安排，通常可采用如下策略。

- （1）按功能分阶段逐步转换。首先确定电子档案管理系统中的一个主要业务功能，如：电子文件和电子档案的收集、整理与归档等功能可率先投入使用，在该功能运行正常后再逐步推广到技术资料的接收、分发、回收及有效性控制等功能。
- （2）按部门分阶段逐步转换。先选择系统中的一个合适部门，在该部门设置终端，获得成功后再扩大到其他部门。首先设置终端的部门可以是业务量较少的部门，这种选择因工作量小相对可靠；也可以是文件处理业务最多

的部门，这种选择效果明显，但风险相对较高。综合考虑二者，选择文书档案管理部门来运行新的电子档案管理系统则更为可行，因为文书档案管理部门对文件和档案的处理比较规范，与档案部门联系比较紧密，出现问题可快速解决。

（3）按机器设备分阶段逐步转换。先从简单的设备开始转换，再推广到整个系统。这种方法只适用于既可用单机又可用联网方式运行的系统，但在文书档案管理部门用得较少。

总之，系统转换的工作量较大，情况十分复杂。据有关统计资料表明，软件系统的故障大部分发生在系统转换阶段，这就要求实施人员要切实做好准备工作，拟定周密的计划，使系统转换不会对整体工作的运转产生不利影响。

## 第七节 系统评价与验收

电子档案管理系统从开发到实施往往会投入一定的资金、人力、物力，但其性能是否达到了预期目标，这是用户和软件开发商双方都十分关心的问题。因此，必须对电子档案管理系统进行评价，形成评价结果。有了评价结果，再进行系统验收，至此，项目实施阶段工作全部完成。

### 一、系统评价

#### 1. 评价目标

（1）检查系统的目标、功能及各项指标是否达到合同或技术协议规定的要求。

（2）检查系统的质量。

（3）检查系统的使用效果。

（4）根据评审和分析的结果，找出系统的薄弱环节，提出整改意见。

## 2. 系统性能评价指标

(1) 系统对用户和业务需求的相对满足程序,即系统是否满足了用户和业务管理的需求,用户对系统的操作和运行结果的满意程度。

(2) 系统的实用性,即考查系统对实际档案工作是否实用。

(3) 系统的开发过程是否规范,包括各开发阶段的工作过程以及文档资料是否齐全规范等。

(4) 系统功能的先进性、有效性和完备性,这是衡量系统质量的关键。

(5) 系统的性能、成本、效益综合比,这是综合衡量系统质量的首选指标,集中反映了一个系统质量的好坏。

(6) 系统运行结果的有效性或可行性,即考查系统运行结果对解决预定的档案工作问题是否有效或可行。

(7) 系统处理结果是否完整,是否全面满足了各级管理者的要求。

(8) 信息资源的利用率,即系统是否最大程度地利用现有的信息资源,并充分发挥了它们在业务管理决策中的作用。

(9) 提供的信息质量如何,即系统所提供信息的准确程度、精确程度。

## 3. 系统运行评价指标

电子档案管理系统在投入运行后,需要不断对其运行状况进行分析评价,并以此作为系统维护、更新和进一步开发的依据。系统的运行评价指标如下。

(1) 预定系统目标的完成情况评价。

对照预定的系统目标和系统建成后的实际情况,评价系统的可维护性、可扩展性、可移植性如何。

(2) 系统运行实用性评价。

评价系统运行是否稳定可靠,系统的安全保密性如何,用户对系统操作、管理、运行状态的满意程度如何,系统对错误操作保护和故障恢复的性能如何等。

## 二、系统验收

电子档案管理系统作为一个信息系统在实施中采用项目制来管理，因此，在实施工作完成后，通过对系统评价后，应进行项目验收，这是项目实施工作是否完成的重要标志。

验收阶段主要工作如下。

### 1. 准备验收资料

起草验收报告，验收还要准备如下资料。

- (1) 系统需求分析报告。
- (2) 合同和技术协议。
- (3) 系统实施方案。
- (4) 系统测试报告。

### 2. 召开验收会

验收会一般由软件供应商和用户双方人员组成，人数在 7~10 人为宜，应参与验收会的人员如下。

- (1) 软件开发及实施项目主管。
- (2) 用户实施主管。
- (3) 用户方档案部门主管。
- (4) 用户方信息技术部门主管。

### 3. 形成验收结论

验收结论是对电子档案管理系统的最终评价，主要内容应包括以下方面。

- (1) 引言。

例如：由 XXX 主持，在 XXXX 召开了 XXXX 电子档案管理系统验收会，会议成立了验收专家组（名单附后），会议听取了系统开发项目组、实施组及用户对系统开发、实施、使用情况汇报，并进行了现场考察，得出如下结论。



(2) 评价条款。

明确开发工作是否规范，系统功能与性能是否达到合同和技术协议要求，是否通过验收等。

(3) 存在问题及解决措施。

(4) 验收组正、副组长签署。

(5) 专家名单及签字表，具体见表 8-1。

表 8-1 专家名单及签字表

序号	姓名	单 位	职务/职称	签字	备 注

## 第九章 标准规范与安全保密体系建设

### 第一节 标准规范建设

企业数字档案馆建设应参考国家相关标准规范并结合企业实际，形成本企业基于数字档案馆的电子档案收、管、存、用等方面的标准规范体系。

#### 一、基础标准规范

基础标准规范是指数字档案馆建设的基础设施建设和各类电子档案管理过程中需使用的标准规范，包括但不限于以下内容。

(1) 名词术语规范。

名词术语规范可不单独制定，采用或参照《电子档案管理基本术语》。

(2) 编码标准规范。

编码标准规范在数字档案馆建设中非常重要，具体包括部门、岗位、保管期限、密级等的编码规则。

(3) 分类规则。

(4) 光、磁载体管理规范。

载体收、存、用、检测规则等。

(5) 安全与保密要求。

(6) 存贮环境标准。

存贮环境标准主要提出机房环境的管理要求。如果不单独建设电子档案管理系统专用机房，则不需要单独制定标准。

(7) 备份技术规范。

(8) 电子档案存储格式规范。

电子档案存储格式规范用于规范电子档案格式注册、迁移、转换等管理过

程,以及通用格式管理,各类电子档案存储格式规范可纳入管理对象标准规范。

(9) 档案信息化设备管理技术规范。

(10) 电子档案管理系统平台功能要求。

(11) 电子档案管理系统与各业务系统接口通用技术规范。

(12) 元数据管理标准。

元数据管理标准要求分批制定各类电子文件元数据标准,包括通用元数据项、捕获节点、专用元数据定义、管理过程等。

(13) 电子文件归档和电子档案管理“四性”检测技术规范。

(14) 电子档案管理系统运维工作规范。

## 二、管理过程标准规范

### 1. 采集工作标准

(1) 电子文件归档登记规范。

(2) 业务系统电子文件归档接口标准。

电子文件归档通用接口规范,适用于所有电子文件归档管理。

(3) 传统载体档案著录规范。

重点在著录项的定义。

(4) 传统载体档案数字化规范。

纸质档案数字化可参照《纸质档案数字化规范》(DA/T 31)制定,缩微胶片数字化参照《缩微胶片数字化技术规范》(DA/T 43)制定。企业如有将档案数字化外包的计划,应将国家有关档案数字化外包安全规范转化为本企业数字档案馆建设标准规范。

### 2. 整理规范

(1) 电子文件归档整理标准。

(2) 传统载体档案计算机辅助整理规范。

### 3. 保存工作标准

(1) 离线备份存储规范。

(2) 电子档案移交与接收规范。

#### 4. 档案资源开发与利用标准

(1) 档案检索利用规范。

规范档案利用审批流程，提出档案利用权限。

(2) 档案利用赋权规范。

主要内容包括企业档案利用赋权原则、权限设置与更改流程等。

(3) 专题数据库建设规范。

#### 5. 安全与保密管理系列规范

应将安全与保密管理要求纳入本企业数字档案馆运行的各项规章制度和标准规范，建立数字档案馆安全与保密职责，制订数字档案馆安全与保密检查、“三员”管理等制度。

#### 6. 岗位职责标准

规范各类人员在数字档案馆运行时的岗位职责，尤其是“三员”的岗位职责必须完备。

### 三、管理对象标准规范

管理对象标准规范包括管理类、产品或业务类、科研类、基本建设类、设备仪器类、会计类、职工类等各类档案在电子档案管理系统中收集、整理、存储、保管和利用等的规范。

#### 1. 管理信息系统电子文件归档管理规范

(1) OA 电子公文归档管理规范。

(2) 财务系统数据归档管理规范。

(3) ERP (MRPII) 数据归档管理规范。

(4) MES 制造执行系统归档管理规范。

(5) 多媒体电子文件归档管理规范。

## 2. 工程技术信息系统电子文件归档管理规范

- (1) PDM 数据归档管理规范。
- (2) CAPP 数据归档管理规范。
- (3) 计算机软件归档管理规范。

## 3. 基础系统电子文件归档管理规范

## 4. 管理对象标准规范的主要内容

管理对象标准规范编制时，可按以下提纲编制。

- (1) 术语与定义。
- (2) 管理原则。
- (3) 管理职责。
- (4) 归档范围。
- (5) 保管期限要求。
- (6) 管理过程要求。

包括收集、鉴定、整理、存储、保管、利用等方面。

- (7) 元数据要求。

# 第二节 安全保密体系建设

## 一、概述

数字档案馆建设与运行必须加强安全管理，遵守相关保密规定，安全设备配置齐全；严格控制信息资源利用范围，制订严密的信息发布审核制度，严格控制数据访问权限；应制订完善的数据接收、移交、销毁等管理规范，制定备份规范。

管理涉密档案的电子档案管理系统，应按照涉密信息系统分级保护管理的要求进行安全设计和建设，并通过有关部门测评后方可投入使用。

管理非涉密档案的电子档案管理系统应根据系统重要程度及其实际安全需

求，参照《计算机信息系统安全保护等级划分准则》（GB/T 17859-1999）确定电子档案管理系统应达到的安全等级，并依照《信息安全技术信息系统安全等级保护基本要求》（GB/T 22239-2008）、《档案信息系统安全保护基本要求》、《档案信息系统安全等级保护定级工作指南》等，开展安全建设和登记管理。

## 二、建设内容

数字档案馆安全保密体系建设应包括物理安全与保密、系统安全与保密、网络安全与保密、应用安全与保密、安全保密管理等内容。

### 1. 物理安全与保密

物理安全包括通信线路的安全、物理设备的安全、机房的安全等。物理安全主要体现在通信线路的可靠性（线路备份、网管软件、传输介质），软硬件设备的安全性（替换设备、拆卸设备、增加设备），设备的备份、防灾害能力、防干扰能力，设备的运行环境（温度、湿度、烟尘），不间断电源保障等。

### 2. 系统安全与保密

系统安全问题来自网络内使用操作系统的安全，如 Windows NT、Windows 2000 等。主要表现在三个方面：一是操作系统自身缺陷带来的不安全因素，主要包括身份认证、访问控制、系统漏洞等；二是对操作系统的安全配置问题；三是病毒对操作系统的威胁。

### 3. 网络安全与保密

主要体现在网络方面的安全性，包括网络身份认证、网络资源的访问控制、数据传输的保密与完整性、远程接入的安全、域名系统的安全、路由系统的安全、入侵检测的手段、网络设施防病毒等。

### 4. 应用安全与保密

主要由提供服务所采用的应用软件和数据的安全性产生，包括 Web 服务、电子邮件系统、DNS 等。此外，还要考虑病毒对系统的威胁。

## 5. 安全与保密管理

安全管理包括安全技术和设备管理、安全管理制度、部门与人员的组织规则等。管理制度化极大地影响着整个网络的安全，严格的安全管理制度、明确的部门安全职责、合理的人员角色配置都可以在很大程度上降低其他层次的安全漏洞。

# 三、安全与保密重点

企业数字档案馆建设是在企业已建成的信息化环境中进行，其物理安全与保密、网络安全与保密、系统安全与保密应由企业信息技术部门统一规划实施，数字档案馆安全保密体系建设重点在应用安全与保密和安全与保密管理上。

## 1. 应用安全与保密

电子档案管理系统本身必须符合安全与保密规定，具有合理的用户赋权和系统日志审核功能。为此，在档案系统的采购中，必须采购符合资质企业的系统，并进行测试，如需管理涉密档案还需进行第三方测试。在赋权方面，规划好系统管理员、安全员、审计员职责，规划各类档案管理人员的权限，规划好各类利用人员权限，规划好各类利用审批流程。

应用的安全与保密还包括档案信息传递必须安全与保密，如数字化加工场所、档案信息存贮，以及提供利用过程的安全与保密等。

## 2. 安全与保密管理

要制定安全保密制度和规范，定期开展安全保密检查，对有关人员进行动态管理等。

## 第十章 企业数字档案馆评价

### 第一节 数字档案馆建设评价问题的提出

由于缺少统一标准，目前数字档案馆建设的效果与水平难以比较与衡量。尽管很多机构都宣称建成了数字档案馆，其建设水平却相差甚远。因为缺少评价指标体系，各企业在对数字档案馆建设目标进行规划时，往往无的放矢，难以确定建设目标和建设重点。因此，构建一个数字档案馆建设的评价指标体系，用于数字档案馆建设评价对于推进我国数字档案馆建设具有重要意义。

#### 一、数字档案馆建设评价

数字档案馆，是集档案资源、实施资源管理的人、硬件、软件、运行机制、制度规范体系等于一体的人机系统，根据系统评价理论，其评价可分为建设评价、性能评价和效果评价。

（1）建设评价，是对数字档案馆建设成果的评价，即规划目标的实现程度，主要是评价建设现状真实值是否达到或超过规划时的期望值。

（2）性能评价，是指对数字档案馆各个组成部分，包括信息化基础设施、信息资源、人员、硬件、软件和各项规章制度等的有机结合，并作为一个整体对使用者所表现出来的技术特性的评价。

（3）绩效评价，是指对数字档案馆运行所获得的收益与成本的比较。绩效包括经济效益和社会效益两个方面。由于数字档案馆运行的效果很难全部量化为可计量的经济效益，因此，绩效评价常常用量化的服务质量评价来代替。



## 二、基于层次分析法的评价指标体系构建思路

国内已有机构设计了打分制的数字档案馆评价方案，并设计了详尽的指标体系，这些方法都可以在企业数字档案馆建设中应用，尤其是中央企业或地区档案行政管理部门制定本集团、本地区的企业数字档案馆建设评价制度时可以采用。打分制本质上是定量评价法，其科学性决定于指标项及指标值（也称权重）的设计。鉴于当前数字档案馆评价制度建设的现状，本书介绍基于层次分析法的数字档案馆建设评价指标体系构建方法，供广大企业档案工作者参考。

层次分析法（Analytic Hierarchy Process, AHP），最早由美国运筹学专家沙特（T.L.Saaty）于 20 世纪 70 年代末提出。

许多具有复杂因素的技术、经济和社会问题，往往很难用定量的模型来分析，因为其中所包含的定性因素很多，而且需要考虑决策者的心理因素、知识经验和决策水平等。而层次分析法则能通过建立所谓判断矩阵的过程，逐步分层将众多的复杂因素和决策者的个人偏好综合起来，进行逻辑思维，然后用定量的形式表示出来，从而使决策者能对复杂问题做到“心中有数”，进而做出正确决策。技术评估本身是一种复杂的决策问题。因此，层次分析法在技术评估中得到广泛应用。

根据层次分析法的原理，用层次分析法确定评价指标体系的思路如下。

第一步：明确问题，建立层次结构。

对于所要解决的问题，首先进行系统分析，明确问题的范围、所包含的因素以及因素之间的定性关系等，然后根据这些初步分析，将各因素分层分组，建立层次结构。

一般层次结构包括最高层、若干中间层和最低层。最高层表示解决问题的总目标，中间层则为各子目标、子准则，最低层则为解决问题的措施、政策或方案。层次结构一般用树形图表示。

具体到数字档案馆的评价，则是先明确数字档案馆评价内容的构成，并用层次的形式表示出来。通过对数字档案馆的层次结构导出数字档案馆的评价因素，由这些评价因素形成一种层次结构。

第二步：构造判断矩阵。

应用所建立的层次结构，可以构造一系列的判断矩阵。首先可以构造最高层目标下属有联系各元素的判断矩阵，然后依次由上而下建立上层某元素与下层有关元素之间的判断矩阵。

第三步：层次单排序。

即对各判断矩阵进行求解，计算出反映上层某元素和下层有关元素重要性次序的权重，并对各判断矩阵进行一致性检验。

第四步：结果分析。

在满足一致性基本问题的前提下，根据层次单排序和层次总排序结果对问题进行定量分析，从而做出选择或决策。

## 第二节 数字档案馆建设评价总体思路

根据数字档案馆层次架构形成的建设内容，构建数字档案馆建设评价指标项如表 10-1。

表 10-1 数字档案馆建设评价指标项构成表

序号	一级评价指标	二级评价指标	指标含义
1	硬件及网络建设	计算机配置率	数字档案馆工作人员计算机配置情况，用每人拥有计算机数表示。
		服务器配置性能	服务器配置是否满足电子档案管理系统运行需求
		网络覆盖率	档案部门计算机接入网络百分比
		档案数字化处理及还原输出设备完备率	档案数字化处理设备、还原输出设备的配备情况，用设备覆盖非数字载体比率表示
2	档案资源数据库建设	档案目录数字化率	馆藏档案目录数字化比率
		档案全文数字化率	馆藏档案全文数字化比率
		电子文件归档率	档案部门每年接收电子文件归档占应接收电子文件的比率
		多媒体档案数字化率	馆藏多媒体档案数字化比率
3	档案管理系统平台建设	采集功能实现度	
		整理功能实现度	
		保管功能实现度	
		统计功能实现度	

续表

序号	一级评价指标	二级评价指标	指 标 含 义
3	档案管理系统平台建设	检索利用功能实现度	
		系统管理功能实现率	
4	运行管理机制建设	岗位建设	数字档案馆建设所需的岗位设置比率
		人员配备	数字档案馆工作人员在专业结构、学历结构和职称结构满足需求的比率
		管理制度	管理制度覆盖相应管理工作环节的比率
5	标准与规范建设	技术性标准规范	标准规范覆盖数字档案馆技术工作的比率
		业务性标准规范	标准规范覆盖数字档案馆业务工作的比率
		管理性标准规范	标准规范覆盖数字档案馆管理工作的比率
6	安全与保密体系建设	数字档案馆系统安全与保密	物理层、网络层、系统层、数据层、应用层五个层面的安全保密建设进展评价，评价方法可借用国家信息系统分级评价方法

### 第三节 企业档案信息化评价指标体系设计

#### 一、评价指标结构的确定

根据前文对数字档案馆评价内容的研究可知，数字档案馆评价内容构成了一个两层架构的层次评价指标结构，其中一级指标 6 项，二级指标 21 项，如图 10-1 所示。

#### 二、权重的确定

根据数字档案馆评价内容和指标项的层次结构，指标权重可用层次分析法确定。

(1) 明确问题，建立层次结构。

将所有的指标用图示的形式表示出来，如图 10-1 所示。

(2) 构造各层次的判断矩阵与层次单排序。

通过求判断矩阵的特征值和特征向量来求得各层次中各分项指标在本层次的重要程度排序，即所谓“权重”。

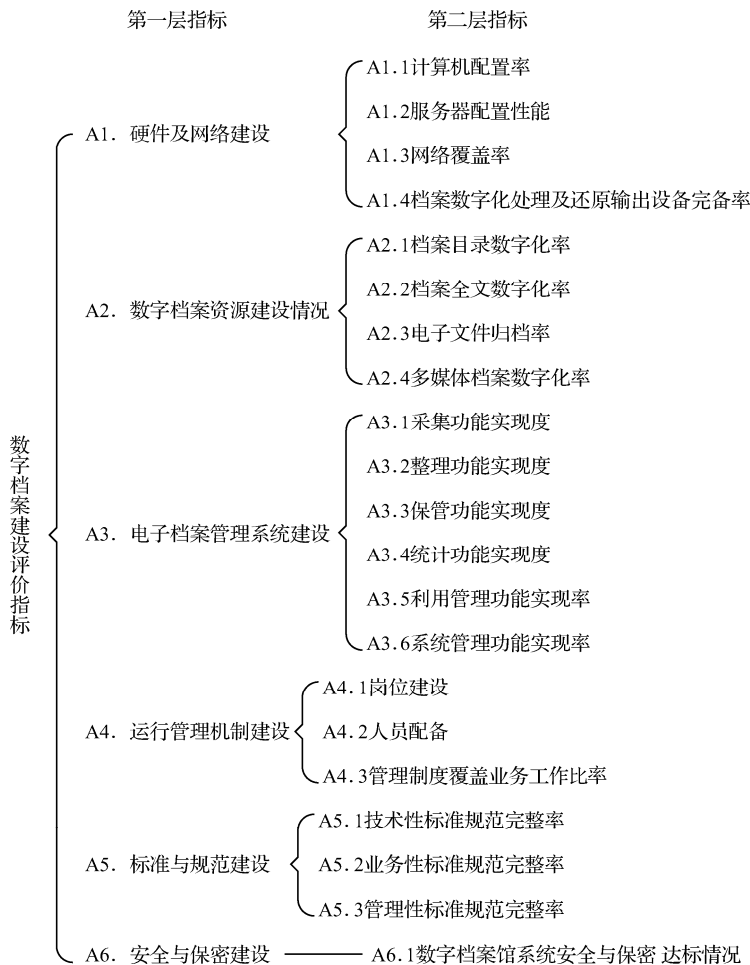


图 10-1 数字档案馆评价指标结构层次图

第一步：构造第一层判断矩阵  $S$ ，见表 10-2。

表 10-2 第一层判断矩阵

$S$	A1	A2	A3	A4	A5	A6
A1	$a_{11}$	$a_{12}$	$a_{13}$	$a_{14}$	$a_{15}$	$a_{16}$
A2	$a_{21}$	$a_{22}$	$a_{23}$	$a_{24}$	$a_{25}$	$a_{26}$
A3	$a_{31}$	$a_{32}$	$a_{33}$	$a_{34}$	$a_{35}$	$a_{36}$
A4	$a_{41}$	$a_{42}$	$a_{43}$	$a_{44}$	$a_{45}$	$a_{46}$
A5	$a_{51}$	$a_{52}$	$a_{53}$	$a_{54}$	$a_{55}$	$a_{56}$
A6	$a_{61}$	$a_{62}$	$a_{63}$	$a_{64}$	$a_{65}$	$a_{66}$

第二步：通过两两比较法确定矩阵中系数的值。这个值可以通过专家调查法来确定。确定后形成系统矩阵  $S$ 。

$$S = \begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} & a_{56} \\ a_{61} & a_{62} & a_{63} & a_{64} & a_{65} & a_{66} \end{vmatrix} \quad \text{式 (10-1)}$$

设权重矩阵  $K=(K_1, K_2, K_3, K_4, K_5, K_6)$

构造矩阵方程： $S \cdot K = 6K$  式 (10-2)

由于  $S$  为已知，利用式 (10-1) 即可求得  $K$ ，也即求得  $K$  的特征向量（唯一解）。

用列和法解此矩阵方程得： $K=(0.19, 0.35, 0.25, 0.06, 0.1, 0.05)$

因此，图 10-1 中第一层次指标权重分别为： $A_1=0.19, A_2=0.35, A_3=0.25, A_4=0.06, A_5=0.05, A_6=0.1$ 。

用同样的方法可求得第二层中  $A_{1.1}$ 、 $A_{1.2}$ 、 $A_{1.3}$ 、 $A_{1.4}$  的权重分别为： $0.21, 0.21, 0.21, 0.37$

用同样的方法可求得第二层中求得  $A_{2.1}$ 、 $A_{2.2}$ 、 $A_{2.3}$ 、 $A_{2.4}$  的权重分别为  $0.172, 0.342, 0.342, 0.144$ ；求得  $A_{3.1}$ 、 $A_{3.2}$ 、 $A_{3.3}$ 、 $A_{3.4}$ 、 $A_{3.5}$ 、 $A_{3.6}$  的权重分别为  $0.2, 0.32, 0.08, 0.16, 0.12, 0.12$ ；求得  $A_{4.1}$ 、 $A_{4.2}$ 、 $A_{4.3}$  的权重分别为  $0.33, 0.34, 0.33$ ；求得  $A_{5.1}$ 、 $A_{5.2}$ 、 $A_{5.3}$  的权重分别为  $0.4, 0.3, 0.3$ ；至此，所有两个层次的权重全部求出。

第三步：把权重放入到体系图中，如图 10-2 所示。

图 10-2 所示的指标体系，为根据现有架构设计，仅提供读者在制定企业数字档案馆评价考核体系时参考，不作为推荐的评价指标。

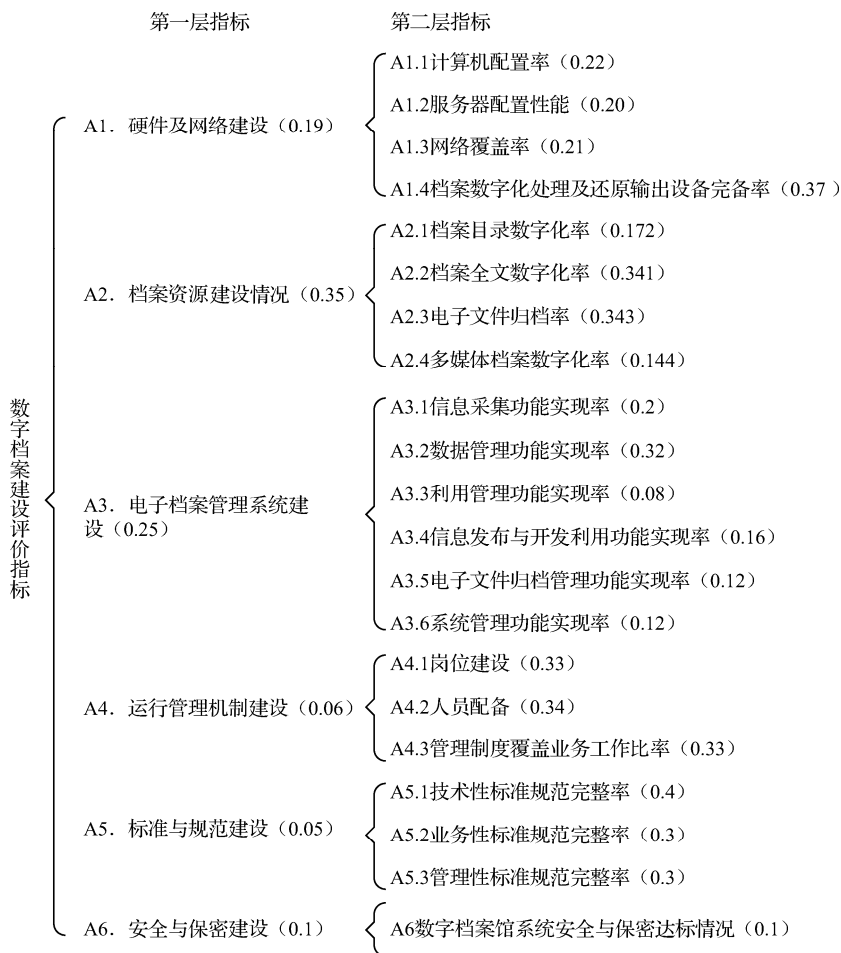


图 10-2 数字档案馆评价指标体系图 (括号内为权重)

# 第十一章 企业数字档案馆与新技术的融合

## 第一节 企业数字档案馆与知识管理

### 一、当前知识管理的突出问题

知识管理作为一场管理的革命，对企业发展具有举足轻重的作用。随着知识经济时代的来临，在未来的全球化竞争中，知识管理能力将成为企业的核心竞争力和可持续发展的关键因素。技术的快速发展、人力资源的市场流动，迫使企业各项活动需要知识管理的支持。随着研究与实践的不断深入，现有知识管理的弊端日益凸显，特别是制造企业，尽管花费巨资购买并实施了知识管理系统，但收效甚微，其问题主要表现在如下方面。

（1）现有的知识管理系统不能对知识进行集成管理，大量基于具体工作任务的知识无法集成到系统中。

（2）受知识来源限制，知识管理对象仍局限于泛化知识，与企业产品和服务高度相关的知识不能进行集成管理，知识利用手段与途径单一。

（3）知识缺乏与数据过剩问题突出。不少企业投入巨资和大量人力物力，构建了以隐性知识为管理对象的知识管理系统。由于隐性知识收集困难，构建的知识管理系统也只是框架，可提供利用的知识内容十分匮乏，逐渐使企业员工对知识管理系统失去兴趣，系统被束之高阁。

利用档案信息作为企业知识管理的知识来源，提出要实现档案信息知识管理，特别是针对以档案信息为知识源的知识表示、知识获取、知识组织、知识利用等关键技术进行研究，找到将档案信息转化为知识并进行管理的技术，为真正实现企业知识管理奠定基础。

## 二、知识管理及其技术

知识管理作为一项活动，由一系列职能活动组成。从信息管理角度，将知识管理职能活动概括为知识的获取、存储、组织、传播、检索、应用等六个方面，其相互关系如图 11-1 所示。

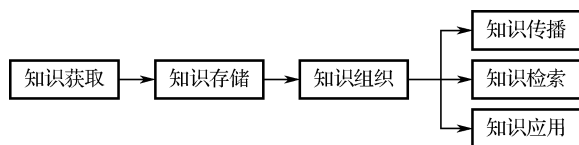


图 11-1 知识管理职能活动图

知识管理作为一项职能活动，它涉及一系列现代先进技术，统称为知识管理的技术。这些技术以现代信息技术为核心，包括数据挖掘、BBS、电子会议、文档管理、搜索引擎、智能检索、E-learning、知识发现、门户系统等，在知识管理系统中具有支撑作用。如果按照知识管理的过程将知识管理涉及的技术进行分类，可将知识管理技术划分为知识获取技术、知识表示技术、知识组织技术、知识利用技术等。

## 三、企业档案信息特点及知识管理难点分析

### 1. 企业档案信息的特点

#### (1) 档案信息蕴含丰富的知识资源

企业档案包括企业在各个时期各项产品研发、生产、销售，企业经营管理、行政管理、生产技术管理和党群工作等过程中形成的所有文件和记录，内容丰富、信息量大。企业档案信息也是企业员工工作过程的原始记录，是企业员工技能与知识转化的结果，是非常丰富的知识资源。档案信息直接产生于本企业的科研、生产和经营管理活动，其中蕴含的信息与企业任务和产品直接相关，特别是蕴含丰富的解决具体问题的实例知识与程序知识。



(2) 数据异构

档案信息以大量数据形式存在，特别是电子档案，形成于不同的业务系统。由于业务系统实施于不同时期，数据异构性十分严重，不仅数据库异构，文档格式、软件架构也异构。据统计，一家制造企业的档案信息，如以数据结构表示的，有 20 余种。

(3) 信息存储格式类型多样

企业档案信息的表现形式是多种多样的，它的数据存储格式有文本、图形、图像文档等多种格式。其中，文本文档格式还可细化为自由文本、半结构化文本、结构化文本、图文混排文本、图表混排文本等多种格式。就企业档案信息来说，关系型数据、网络数据、链式数据等结构化数据所占比例较小，而格式多样的非结构化数据占绝大部分，经统计，仅企业的文书档案，信息存储格式就达 50 余种。

(4) 知识信息与非知识信息混杂

档案信息是从长远利用的角度经过档案人员甄别并保存的原始记录信息，其中既有可转化为知识的有用信息，即知识信息，也有无知识价值的信息，即非知识信息。这些非知识信息作为追溯与凭证的需要而留存。

2. 档案信息作用于知识管理的机制与难点

知识管理活动一般以文献和专家的隐性知识为知识源，结合知识管理活动过程可整体表示知识管理流程如图 11-2 所示。档案信息在知识管理中主要起着提供知识源的作用，其作用与文献信息、专家的隐性知识源类似。以档案信息为知识源的知识管理流程如图 11-3 所示。

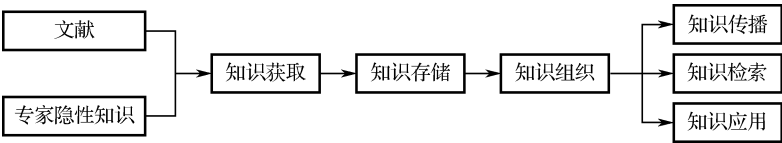


图 11-2 传统知识管理活动过程图

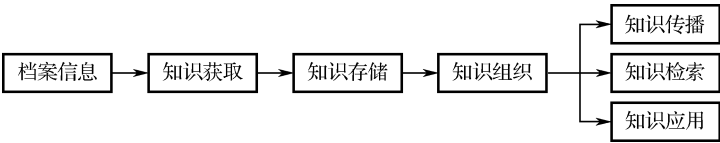


图 11-3 档案信息的知识管理活动过程图

比较图 11-2、图 11-3，两者的不同点在于知识源不同，前者是以文献和专家隐性知识为知识源，后者是以档案信息为知识源。由于知识源的变化，知识管理过程中涉及的技术也必须随着知识源的改变而改变，以档案信息为知识源的知识管理流程需要采用可处理档案信息的知识管理技术加以实现。档案信息知识管理的难点主要表现在以下方面。

#### （1）在知识表示上的难点

知识表示过程就是把知识编码成某种数据结构的过程。知识表示技术的选择取决于需要表示知识的所属领域、知识的利用方式以及利用目的，同时也取决于需表示的知识源。从知识源来考虑，知识表示应选择更加接近知识源的表示形式，从而减少转换过程，并尽可能表示所有的知识。由于档案信息在内容与结构上的复杂性特点，与文献知识源和专家隐性知识源相比，档案信息知识源的表示方法要有更强的表示能力，除要求知识表示不影响存储和利用外，表示方法还应兼容多种格式的档案信息知识源，并且尽可能不改变档案信息的原始形式。

#### （2）在知识获取过程中的难点

知识获取是指将未经组织的文档、数据等（显性知识）和存在于人脑的专家技能（隐性知识）转化为可复用、可检索形式的知识。由于档案信息数据异构、存储格式多样、结构化低、数据冗余等方面的特点，档案信息通过获取转化为知识存在较大的难点。主要表现为：档案信息表现形式、存储形式多样且数据异构，使档案信息获取为知识很难用单一的方法来完成；档案信息在数据表现形式上的非结构化与知识的结构化要求存在较大差距，获取过程要对大量非结构化、半结构化的档案信息进行结构分析，给知识获取带来较大的挑战；档案信息的高冗余使知识内容去重工作存在较大困难。

#### （3）在知识组织和知识利用上的难点

知识组织包括知识存储与呈现。由于知识存储与知识利用都直接受知识表示的影响，因此，在基于档案信息的知识管理过程中，存储和利用同样存在难点。知识表示的方式直接决定知识存储的方式。文献信息和专家知识由于其结构化好，内容单一，知识特点明显，在管理上相对容易。与文献信息和专家头脑中的隐性知识相比，档案信息具有内容丰富、表现与存储形式多样、数据异构严重的特点，在知识存储和呈现过程中必须兼顾多种数据格式，采用多种呈现方式，并且不损失知识利用效率。由于知识组织上的难点，需采用更加有效的存储和呈现方式，这将给知识利用带来困难。

## 四、知识管理技术的运用

### 1. 档案信息的知识获取

在数字档案馆建设中，知识管理技术应用的第二种方案是将数字档案馆中的档案信息作为企业知识库的知识源。这要求企业已经开展了知识管理，或建有知识库。

档案信息作为企业知识库的知识源，其重点在于将档案信息以知识获取的方式转化为知识，也即档案信息的知识获取。

根据档案信息的特点，档案信息知识管理的知识获取方法必须采用基于规则的知识自动获取技术，该技术能针对不同的知识或来源，选择相应的规则进行知识的自动获取，具体原理如图 11-4 所示。

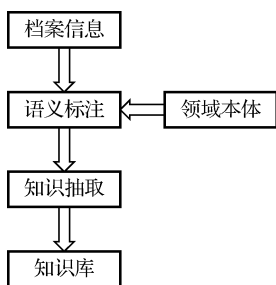


图 11-4 档案信息知识获取过程图

语义标注是读取档案信息用语义方式标注出来。知识抽取是通过一定的算法从档案信息中抽取出知识，即运用信息抽取算法，把具有知识属性的内容提取出来。针对档案信息复杂性的特点，需对现有的算法进行改进，采用基于多层规则的知识抽取技术，将原来针对词句的一层规则改进为分别针对文档和段落的三级规则，这样获取效率更高。由于企业档案信息中含有大量图形，包括二维和三维的图形，所以基于档案信息知识管理的知识获取方法还必须能够获取图形知识。图形文档一直被认为是不可进行语义标注的文档，但如采用 STEP 规范，将图形文档转化为 STEP 结构的数据，然后在此基础上进行语义化，就较好地解决了此类知识的语义标注问题。

## 2. 档案信息的知识组织方案

数字档案馆建设应用知识管理技术的第三种方法便是将电子档案库进行知识组织，将档案信息直接转化成知识，并以知识库的形式组织并呈现出来，但这种组织和呈现会改变档案信息原有的组织方式。

知识组织包括存储和呈现两个方面。在存储方面，基于档案信息的知识存储必须考虑知识的多种形态、多种文档，且方便利用。在知识呈现方面，一方面要反映知识的来源，另一方面要反映知识的逻辑关系。基于上述要求，档案信息知识组织采用基于本体的层次知识存储法，如图 11-5 所示，该图将来自档案信息的产生层、存储层、本体层、主题层四层来呈现。产生层表示知识产生的任务来源，保证企业每一活动过程和产品每一阶段产生的知识得到管理；存储层描述企业知识的具体存储形式，其实是档案存储；本体层用语义形式描述企业知识，它描述了企业知识的概念划分及其相互关系，实质是用本质来描述企业档案，是对档案数据进行知识标引；主题层描述了企业知识所属的主题，为知识利用提供基础。基于本体的知识层次组织法能较好地解决企业知识内容多样性、分布性、异构性等问题。

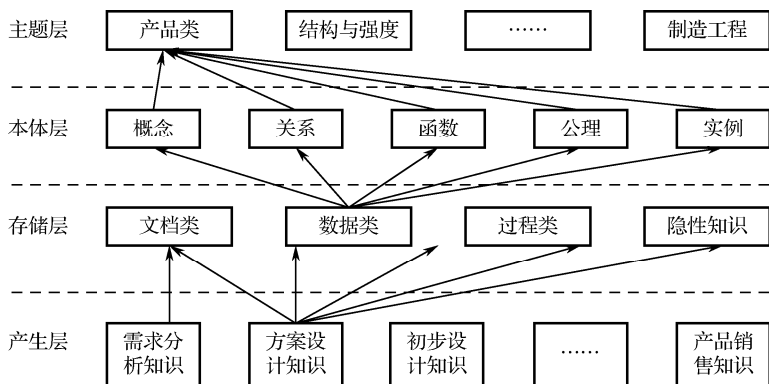


图 11-5 档案信息的知识存储模型图

在实践中，将电子档案数据库中的档案信息用知识标签标记出来，以方便检索是一种比较实用的方法。

## 3. 档案信息的知识检索利用方案

知识检索取决于是否得到利用，能否将知识及时准确并且方便地传递给企

业的各项活动。在知识管理的任何一个领域都必须涉及知识检索。知识的检索方式决定于应用目的，一直以来知识利用目的是知识的转移，因此，知识共享是其主要目的，增强知识共享的方式主要有知识呈现、知识社区、知识检索和知识推送等方式。而知识工程从人工智能发展而来，主要是将知识运用于具体的任务系统和基于知识的工程，主要应用方式有推理、重用及知识发现。在这些知识应用方式中，检索在知识应用中占有重要的地位，它使应用以最高效的方式找到知识。在企业档案信息知识管理中，其知识来源于企业活动的全过程，服务对象也是企业活动全过程，因此，对知识检索提出了更高的需求。在利用方式上，应综合采用多种方式进行，包括知识地图、知识社区、知识检索和基于知识工程的知识推理、知识重用等的利用方式。由于在知识地图、知识社区等方面已有成熟的理论与应用，基于本体的知识检索、知识推理和知识重用的有关应用技术是档案信息知识利用的重点，由于来自于档案信息的企业每一项工作的反映，会大量以实例知识的形式表示出来，因此知识重用是比较有效的利用方式，并会在实例的检索上进行算法的改进与创新。

在档案检索方面，当前应用知识管理技术的一个可行的方法，便是对档案增加知识标签，并形成与传统方法不一样的档案呈现方法。如某企业将档案增加标签后，以一种方便档案利用的方式呈现，见表 11-1。

表 11-1 档案信息的呈现方式

文书档案	科技档案	会计档案	技术档案	会议档案	特种载体档案	专题档案
集团发文	销售合同	凭证档案	甲产品	董事会	照片	大事记
党委发文	法务合同	账本档案	乙产品	监事会	各种证照	战略合作协议
集团收文	财务合同	报表档案	丙产品	办公会	视频会议	信息简报
集团发函	建设项目	ERP 数据	丁产品	会议纪要	声像	内部期刊
集团签报	采购项目			计划会议	荣誉	出访档案
财务文件	审计档案			财务会议	印章	来访档案
投资文件	巡视档案			人事会议		会见活动
销售文件	纪检档案			投资会议		来信来访
	房产物业					评先文件
	安全事故					职称评聘
						机构沿革
						干部任免
						内网信息

上表是某企业检索界面的档案呈现图，它的特点是，档案信息的分类呈现不是按照档案人员的习惯，而是按照利用者的需求和理解，非常有利于档案的检索和利用。尽管该分类方法既存在交叉，也存在重复，但毕竟不是用来编号的分类，仅是用来呈现档案信息和检索档案入口之用，体现了利用优先的数字档案馆建设思想。要实现上述做法，就需对每条档案目录进行标签定义，这就是利用知识管理的理念与技术提升数字档案馆的应用成效。

## 第二节 企业数字档案馆与大数据

### 一、大数据及其特点

“大数据”在麦肯锡全球研究所给出的定义是：一种规模达到在获取、存储、管理、分析方面大大超出了传统数据库软件工具能力范围的数据集合，它的特征具有以下几个方面。

#### 1. 数据量大

大数据的第一个特征是数据量大，包括采集、存储和计算的量都非常大。大数据的起始计量单位至少是 P（1P=1024T）、E（1E=1024P）或 Z（1Z=1024E）级。

#### 2. 类型繁多

第二个特征是种类和来源多样化。包括结构化、半结构化和非结构化数据，具体表现为网络日志、音频、视频、图片、地理位置信息等，多类型的数据对数据的处理能力提出了更高的要求。

#### 3. 价值密度低

第三个特征是数据价值密度相对较低，或者说是“浪里淘沙”却又“弥足珍贵”。随着互联网以及物联网的广泛应用，信息感知无处不在，海量信息，但价值密度较低，如何结合业务逻辑并通过强大的机器算法来挖掘数据价值，是

大数据时代急需解决的问题。

#### 4. 速度快时效高

第四个特征是数据增长速度快，处理速度也快，时效性要求较高。比如搜索引擎要求几分钟前的新闻能够被用户查询到，个性化推荐算法尽可能要求实时完成推荐。这也是大数据区别于传统数据挖掘的显著特征。

## 二、大数据与档案工作的融合

大数据管理本质上也是信息处理工作，它和档案管理的本质是一样的。

#### 1. 大数据具有与电子档案相同的管理流程

根据大数据原理，大数据的管理流程是采集、存储、处理、挖掘利用等前后相继的几个环节，这些环节与电子档案管理流程如出一辙，二者均有相同的收（采）集流程，相同的处理、存储流程，相同的利用流程等，因此，大数据和电子档案的管理流程相同。

#### 2. 大数据具有与电子档案相同的来源

企业可利用的大数据，主要来源企业各项活动及各业务系统运行产生的数据。电子档案是企业各项活动的记录，也正是来源于企业的各项活动。

#### 3. 大数据具有与电子档案相同的特征

大数据具有数据量大、种类多、异构等特点，而这些也正是电子档案的典型特征。电子档案由企业的大部分原始记录形成，数据量大。电子档案不仅在数据存储格式方面种类繁多，处理种类繁多的数据同样是信息时代档案工作面临的一大挑战。

#### 4. 大数据具有与电子档案相同的管理目的

大数据管理的目的，是要将经过分析处理后挖掘得到的数据结果应用于管

理决策、战略规划等，这以电子档案发挥参考作用服务于企业的决策、规划等的管理目的是一样的。当然，参考作用不是电子档案管理的唯一目的，但就参考作用的目的来看，二者也是一致的。

因此，大数据管理和电子档案管理可以融合，互相促进，相得益彰。尤其是数字档案馆的建设，为大数据应用储存了大量数据，同时运用大数据技术可以有效提升数字档案馆在电子档案信息开发利用方面的成效。

### 三、大数据技术在数字档案馆中的应用

在大数据背景下，因海量无限、包罗万象的数据存在，让许多看似毫不相干的现象之间发生一定的关联，使人们能够更简捷、更清晰地认知事物和把握局势。大数据的巨大潜能与作用，目前难以进行估量，但揭示事物的相关关系无疑是其真正的价值所在。

大数据分析及挖掘技术是大数据应用的关键。数据挖掘就是从大量、不完全、有噪声、模糊、随机的实际应用数据中，提取隐含其中的、人们事先未知的但又具有潜在利用价值的信息和知识的过程。

以数字档案馆建设为契机，在数字档案馆建设中融合大数据技术，是提升数字档案馆建设水平的有效途径，二者融合的方法主要有以下几种。

#### 1. 扩大电子文件收集范围

在大数据背景下，让许多看似毫不相干的现象之间发生一定的关联，使人们能够更简捷、更清晰地认知事物和把握局势。大数据的这一特点使档案工作中曾被认为是大量重复产生的低价值或无价值的电子文件变得有价值。因此，在数字档案馆建设中需要考虑到这种情况，对现有的电子文件归档范围进行相应的研究，使之更加符合大数据的需要。同时，数字档案馆建设通过传统载体档案数字化，可以将信息技术应用前的数据纳入大数据应用范围，使原来记录于传统载体上的不可能纳入大数据应用的记录加入进来，进一步扩大了大数据应用的范围和时间跨度，使大数据分析模型得到的结论更加准确。



## 2. 采用有利于应用大数据技术的电子档案存储方式

当前采用的电子档案存储方式是“元数据+独立文档”形式,在这一方式下,许多文档是无法提取信息的。为了应用大数据技术,数字档案馆建设可以在这一方式的基础上,探索有利于应用大数据技术的信息存储和组织方式。

## 3. 采用大数据技术建设专题数据库

建设专题数据库是数字档案馆发挥效用的有效措施,也是数字档案馆特有的档案利用方式。前期,有企业摸索出了一些经验,但这些经验还比较初级。在大数据时代,应用大数据的分析模型可以从更符合企业业务需要的角度建设专题数据库,使专题数据库发挥更大作用。

# 第三节 数字档案馆与云服务

## 一、云服务概述

### 1. 云服务及架构

云服务是指将计算机软件、硬件、数据等计算资源,从本地迁移至云端,用户随时随地通过成本较低的客户端连接到处于云端的功能强大的软件、硬件和数据资源,其主要目的是实现客户端计算成本的最小化和云端利益的最大化。

云服务包括以下几个层次的服务架构。

#### (1) IaaS

IaaS (Infrastructure-as-a- Service, 基础设施即服务)。IaaS 是消费者通过 Internet 可以从完善的计算机基础设施获得服务。IaaS 是把数据中心、基础设施等硬件资源通过 Web 分配给用户的商业模式。

#### (2) PaaS

PaaS (Platform-as-a-Service, 平台即服务)。PaaS 是指将软件研发的平台作为一种服务,以 SaaS 的模式提供给用户。

### （3）SaaS

SaaS（Software-as-a-Service，软件即服务）。SaaS 是一种通过 Internet 提供软件的模式，用户无需购买软件，而是向提供商租用基于 Web 的软件，来管理企业经营活动。SaaS 模式大大降低了软件成本，尤其是大型软件的使用成本，并且由于软件是托管在服务商的服务器上，减少了客户的管理维护成本，可靠性也更高。

## 2. 云模式

云模式主要有以下几种分类。

### （1）公有云

公有云通常指第三方提供商为用户提供的能够使用的云。公有云一般可通过 Internet 使用，可能是免费或者成本低廉。

### （2）私有云

私有云是一个组织单独使用而构建，提供对数据、安全性和服务质量的最有效控制。私有云可由 IT 部门或云平台业务提供商搭建。企业可以在搭建的云平台基础上部署自己的网络或应用服务。私有云可部署在企业的数据中心，也可统一部署在云平台业务提供商的机房。

### （3）混合云

混合云是公有云和私有云两种服务方式的结合。由于安全和控制原因，并非所有的企业信息都能放置在公有云上，因此大部分已经应用云计算的企业将会使用混合云模式。

### （4）移动云

移动云是把虚拟化技术应用于手机和平板电脑。适用于移动设备终端（手机或平板电脑）使用企业应用系统资源，它是云计算移动虚拟化中非常重要的一部分。

### （5）行业云

行业云是由行业内或某个区域内起主导作用或者掌握关键资源的组织建立和维护，以公开或者半公开的方式向行业内部或者相关组织和公众提供有偿或无偿服务的云服务。

### 3. 云服务的特点

#### (1) 超大规模

云服务的服务器数量、存储能力与计算能力等均规模超大。

#### (2) 虚拟化

云计算支持用户在任意位置使用各种终端获取应用服务。

#### (3) 通用性

云计算不针对特定的应用，同一个“云”可以同时支撑不同的应用使用。

#### (4) 轻松共享数据

云计算可以轻松实现不同设备间的数据与应用共享。

#### (5) 高可扩展性

“云”的规模可以动态伸缩，满足应用和用户规模增长的需要。

#### (6) 按需服务

“云”是一个庞大的资源池，可按需购买服务。

## 二、数字档案馆云服务类型

鉴于云计算带来的便利性以及不同类型的服务，它逐渐应用于电子档案管理中。根据云服务的用户角色不同，电子档案管理应用的云服务类型主要有以下三种。

(1) 基础设施服务模式，主要为数字档案馆提供基础设施，将数字档案馆中的网络资源、存储资源和服务器资源等基础设施作为一种服务资源向外发布。

(2) 平台服务模式，主要为数字档案馆提供基础设施层的平台资源服务，该服务为电子档案管理系统提供运行所必需的软件平台资源，包括软件开发环境，各类数据存储环境、中间件等，数字档案馆可以基于平台服务所提供的服务接口，直接在该平台上进行开发或部署电子档案管理应用。

(3) 软件服务模式，通过对数字档案馆应用进行整合，以服务的形式向用户发布各种电子档案管理应用，终端用户不仅可以通过数字档案馆门户网站对电子档案资源加以使用，还可以直接通过网络或者接口使用电子档案应用服务。

### 三、数字档案馆云服务机遇和面临的挑战

云服务是对信息与通信技术的创新，这也改变了电子档案形成与管理的环境，内容与形式也随之受影响。一方面，业务活动在云环境中开展，使得原生电子文件在云环境中形成并管理，从技术、管理、司法等多方面提出挑战，这需要依据电子文件管理的要求与云环境的特性相结合形成相对完善的管理方案。另一方面，云服务为电子档案提供新的存储与管理环境，非云环境中形成的电子档案可迁移至云环境中，使用各类硬件、平台、基础设施、软件等云服务，这也对迁移乃至之后的管理提出挑战。以美国和澳大利亚电子档案管理对云服务的应用来看，它对数字档案馆的挑战呈现于方方面面，如云应用可能缺少实施档案处置计划的能力、云服务提供商对档案保存要求了解有限、不同的云机构缺乏正式的技术标准来知道数据在云环境中如何存储与管理、难以依据档案管理要求移除档案或是将档案迁移至另一环境、云服务提供商的业务运营可能消失或终止，等等。

而在中国，云服务不仅逐渐开始被应用于业务活动中，也被应用于数字档案馆的建设。如美、澳等国一样，中国同样注重云服务对电子档案真实性、完整性、可用性以及安全性等方面的影响，尤其是云服务相关法律与政策不尽完善，且电子档案管理对云服务只是在有限应用的情况下，如何在采用云服务时保障电子文件的安全与保密是当前提出的首要挑战。例如，利用安全、销毁安全、移交安全等被视作需要重点攻克的技术与管理难题。因而，有必要从安全与保密方面调查与梳理可能存在的风险，为云服务的可靠应用提供基础。

### 四、基于云服务的数字档案馆架构及实现

#### 1. 基于云服务的数字档案馆架构

根据云服务架构，基于云服务的数字档案馆架构见图 11-6。

#### 2. 基于云服务的数字档案馆实现方式

数字档案馆在云服务中的实施其实就是数字档案馆各部分功能在云中的实

施，按照数字档案馆的功能模块划分，各功能模块的实现方式如下。

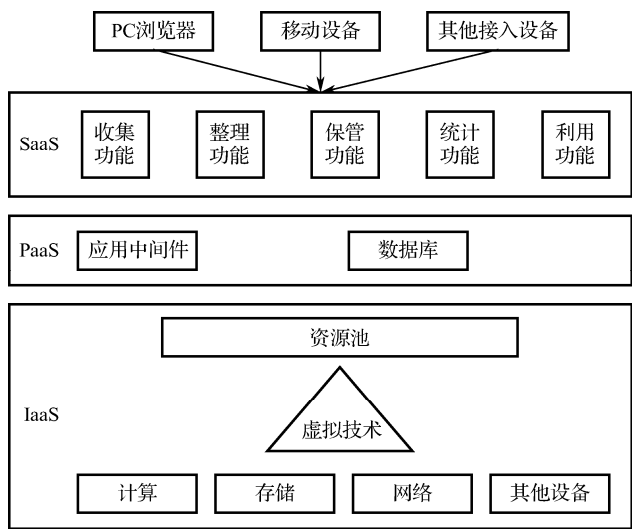


图 11-6 基于云服务的数字档案馆架构图

（1）收集功能实现

通过可扩展接口在线或离线接收来自各信息系统的电子文件和其他数字资源，并采用安全的传输方法保证数据传输安全。

（2）管理功能实现

基于云计算的 SaaS（软件即服务）提供应用程序开发、数据库、应用服务器等服务，定制开发数字档案馆所需应用程序，以满足数字档案管理方面的功能需求。还要通过 PaaS 实现云服务的管理和监控，以监控应用运行的情况和具体数值，监控应用的运行状态。

（3）存储功能实现

利用云存储技术，提供数据存储、访问功能。云服务方建立相应的数据备份系统、容灾系统，定期对存储设备进行状态监控、维护、软硬件更新和升级，对数据移动进行端对端加密。通过云存储可以解决海量数据的储存，实现随着档案存储节点的增加而无限地扩容。通过档案云存储满足数字档案长期保存的要求，实现数据存放及备份、内容搜寻、数据大量读取、多用户共享数据等功能。

（4）利用功能实现

通过定制开发基于 SaaS（软件即服务）的应用，实现档案信息资源发布、

检索、开发、复制等所需的技术工具，由 SaaS 平台服务方负责维护、管理这些工具。而档案利用用户通过接入云服务便可使用云中数字档案馆提供的各项服务，无需安装、升级客户端系统。云服务实现档案利用以设备为中心向以信息为中心的转变，并实现接入端工具的轻薄化、可移动化，降低了用户端设备的需求。

## 第十二章 数字档案馆建设案例

### 第一节 某央企集团公司数字档案馆建设案例

#### 一、工作步骤

某集团公司数字档案馆建设工作分为以下五个步骤进行。

##### 1. 准备阶段

###### (1) 调研学习

组织有关人员去国内已开展数字档案馆建设的单位参观交流。

(2) 组织制定《某集团企事业单位数字档案馆建设方案》，明确建设内容与任务。

(3) 组织制定《某集团企事业单位数字档案馆认定办法》，明确某集团企事业单位数字档案馆的认定与分级标准，明确对各单位数字档案馆建设工作及水平的组织和评价办法，规范工作流程，明确所需的文件材料。

###### (4) 选定试点单位，组织试点工作。

按照试点单位业务类型全覆盖的原则，选定10家企事业单位开展试点工作。试点单位通过如下程序确定：总部拟定试点条件；企事业单位提出试点申请，申报试点方案；总部组织专家对试点方案进行评审，形成评审意见；总部根据评审意见确定试点单位名单。

##### 2. 动员阶段

(1) 召开全集团数字档案馆建设工作动员会暨现场会，会议主要内容如下。

① 全面启动数字档案馆建设工作。

② 试点单位介绍开展数字档案馆建设工作经验，并现场参观学习。

③ 讨论确定试点企事业单位数字档案馆验收计划安排。

(2) 组织宣贯《某集团企事业单位数字档案馆建设方案》和《某集团企事业单位数字档案馆认定办法》。

(3) 各单位制定本单位数字化档案馆建设方案和实施计划，报总部备案。

### 3. 组织实施阶段

(1) 各单位按照本单位数字档案馆建设方案和实施计划具体组织建设工作。

(2) 建设过程中的咨询指导。

总部不定期汇编有关数字档案馆建设的论文、技术报告、科研成果等印发至各单位，供各单位在数字档案馆建设中参考。总部组建数字档案馆建设咨询指导专家组，深入各单位进行指导。

### 4. 考核认定阶段

(1) 各单位按照《某集团企事业单位数字档案馆认定办法》进行自检。

(2) 自检合格后提出数字档案馆认定申请及认定等级。各单位根据建设情况，可一次性申请一级数字档案馆的认定，也可先申请三级或二级数字档案馆认定，之后再申请一级认定。

(3) 按照《某集团公司企事业单位数字档案馆考核认定办法》组织考核认定。

(4) 下达通过考核认定的事业单位数字档案馆名单及认定等级，并授牌。

### 5. 总结提高阶段

(1) 组织交流数字档案馆建设工作经验，对先进单位和先进个人进行表彰。

(2) 对已认定为相应等级的数字档案馆，每两年组织一次复查，或对已认定为二、三级的数字档案馆进行升级认定。

## 二、主要内容

### 1. 开展数字档案馆建设的宣传动员工作

通过各种会议、活动等场合宣讲数字档案馆建设的重要性、必要性和推进



思路等，为数字档案馆建设创造良好的工作氛围。

## 2. 开展数字档案馆建设的顶层设计工作

编制《某集团公司企事业单位数字档案馆建设方案》和《某集团公司企事业单位数字档案馆等级认定管理办法》。编制前者的目的是为企事业单位数字档案馆建设提出一个总体的指导框架，避免建设中各自为政。编制后者的目的是对各企事业单位数字档案馆建设成果有一个统一的评价标准。

两个文件的编制过程如下。

(1) 由总部组织成立文件编制工作小组，小组成员 21 人，召开工作小组全体成员会议，确定文件主体框架和编制进度。

(2) 编制小组完成文件初稿的编写、修改等工作。

(3) 将初稿发送行业内的有关专家进行评审，并反馈评审意见。参与评审的有本集团档案信息化专家组成员、高校教授、国家档案局领导等十多位专家。

(4) 按照专家评审意见对初稿进行修改完善后，将文稿发送至各企事业单位征求意见并反馈。

(5) 按照各单位反馈意见对文稿进一步修改完善。

## 3. 开展数字档案馆建设的关键技术研究工作

数字档案馆建设尚存部分有待解决的关键技术。自集团启动数字档案馆建设工作起，总部组织了有关企事业单位开展如下数字档案馆建设的关键技术研究。

- (1) 数字档案馆总体架构与解决方案研究。
- (2) 数字档案馆应用平台和网络建设研究。
- (3) 档案资源库建设研究。
- (4) 数字档案馆服务模式创新建设研究。
- (5) 多媒体档案管理系统应用研究。
- (6) 档案信息网站的规划与建设研究。
- (7) 电子档案接收、管理技术方案的研究。
- (8) 网络系统中档案信息资源的安全保密技术应用研究。
- (9) PDM 实施与电子档案管理技术研究。

这些研究基本涵盖了数字档案馆建设有关的关键技术。经过一年多的研究，攻关工作全部完成，形成了研究报告，并通过了专家组的技术鉴定。关键技术研究工作的完成，为数字档案馆建设工作奠定了坚实的技术基础。

#### 4. 开展数字档案馆建设的相关制度规范编制工作

从一开始，总部档案部门就组织有关人员开展数字档案馆标准规范的研究，近期又组织档案信息化专家组成员开展有关标准规范的研究，数字档案馆标准规范体系正日臻完善。

### 三、主要成果

经过两年多的建设，该集团公司数字档案馆建设取得了丰硕的成果。试点工作启动后 1 年半时间里，第 1 家试点单位——某研究所数字档案馆建设即通过了集团公司企业事业单位一级数字档案馆验收，后续 1 年中，又有近 10 家企事业单位通过相应等级的数字档案馆验收。其中，通过一级数字档案馆验收的有 6 家，二级的有 2 家，三级的有 1 家。试点结束后，该集团公司适时全面启动了数字档案馆建设工作。经过 5 年的建设，有近 60 家企事业单位的数字档案馆通过了验收。

该集团公司通过推进数字档案馆建设，使全集团档案信息化水平得到了大幅度提高，取得了令人满意的效果。一是解决了业务系统形成电子文件归档的问题，使 PDM、ERP 等核心业务系统数据实现了归档；二是传统载体档案数字化率进一步提高，提升了档案信息服务水平；三是形成了一批行业档案信息化建设标准规范，总数有几十份；四是培养了既有档案管理技能又有信息技术技能的专业人才队伍；五是档案部门在企业中的地位大幅提高。

## 第二节 国泰君安证券公司数字档案室建设案例

国泰君安证券股份有限公司（以下简称“国泰君安”）根据《企业数字档案馆（室）建设指南》和《数字档案室建设评价指标》等标准规范，在原档案管

理系统基础上，优化收集、管理、存储、利用等功能，开发电子文件归档接口，实现传统载体档案与电子档案的全生命周期管理；同时引用智能技术深入挖掘档案资源价值，实现“知识化”档案价值利用，搭建公司全系统内统一共享利用的档案信息化管理平台，全面提升档案管理现代化水平，充分发挥支持功能，为公司各项业务的良性发展提供坚实基础。

## 一、建设背景

在当前互联网金融时代，随着证券业务办理模式的转变和证券公司创新业务的不断发展，电子文件（档案）越来越多，如何对这些非传统载体档案进行管理，是企业档案工作面临创新转型的主要问题。

企业的综合治理能力现代化离不开档案数字化、管理信息化和利用网络化，而企业数字档案室是推进企业治理体系和治理能力现代化的重要工具。通过企业数字档案室建设将存量纸质为主的档案数字化，转换为数字档案信息，将增量的电子文件归档直接进行电子化管理，同时，对海量档案信息数据提供更智能化、便捷化、快速化的全文级检索查询利用，这是企业现代信息化发展的客观要求，是企事业单位档案信息化建设的重要内容，是提高本企业档案管理水平的必然选择，也是未来企业档案工作的努力方向和发展目标，所以推进这项工作刻不容缓。

配合公司创新业务不断发展，特别是随着信息化程度的不断提高，抢占企业现代化、信息化管理前沿，将是企业未来发展的核心竞争力。因此，建设与公司业务信息化相匹配的，行业内先进的企业数字档案室，推动企业档案管理迈入数字时代，落实公司数字化发展战略，以更好地为公司未来在网络化、信息化、智能化时代不断发展壮大提供服务和支持。

## 二、建设目标

依据《企业数字档案馆（室）建设指南》和《数字档案室建设评价指标》等标准规范，以及电子档案信息资源规范要求和业务协同应用原则，在原档案管理系统基础上，搭建公司全系统内统一共享利用的档案信息化管理平台，优化收、管、存、用等功能，开发电子文件归档接口，既扩大传统载体档案覆盖

范围，又尽可能多地将公司业务系统电子文件直接在线归档，实现传统载体档案与电子档案的全生命周期管理；同时引用大数据分析、人脸识别、签名比对、自然语音识别等智能技术深入挖掘档案资源价值，实现“知识化”档案价值利用，全面提升档案管理现代化水平，充分发挥支持功能，为公司各项业务的良性发展提供坚实基础，并努力成为行业内档案管理的龙头标杆企业。具体来说，企业数字档案室系统建设目标如图 12-1 所示。

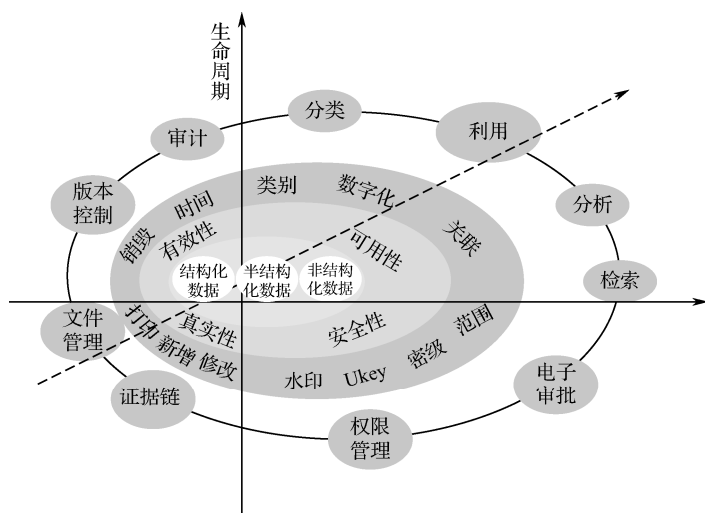


图 12-1 数字档案室建设目标

## 1. 基础设施建设目标

主机房配置科学合理，配备能满足数字档案资源管理需要的网络基础设施，配备能满足工作需要的专业服务器，配备数字档案资源在线存储设备；配备近线或离线备份系统；配备必要的正版基础软件；参照信息系统安全等级保护有关要求为数字档案室应用系统建立安全保障体系；配备能满足工作需要的终端及辅助设备。

## 2. 数字档案资源建设目标

将传统载体档案及证券业务、会计、声像类等各类电子档案纳入企业数字档案室系统集中统一管理。在现有数字档案资源建设的基础上，建立统一接口平台，进一步规范扩展公司各业务系统电子化归档连接，实现与集中营运平台、

大投行项目管理平台、统一风控平台、资产托管综合管理平台、固定收益审批事务平台等对接，实现相应电子文件电子化归档及时，齐全完整，整理规范，满足“四性”保障要求，提高公司现代化、信息化水平。

### 3. 应用系统建设目标

企业数字档案室要具有各业务系统电子文件归档功能、档案门类管理、配置信息管理、接收采集功能、分类编目功能、检索利用功能、鉴定统计功能、监督指导功能、系统管理功能等，支持 7×24 小时运行，系统平均响应时间不超过 5 秒，10 用户并发访问时，系统运行正常，服务器 CPU 使用率<70%，内存使用率<80%。

### 4. 档案信息开发利用目标

支持专题库、借阅利用、档案信息开发、档案编研等档案信息开发利用功能。同时，优化人脸识别、签名比对、语音识别在档案编研及风控、合规、审计中的功能应用。

### 5. 制度规范建设目标

以国家相关标准规范为准绳，结合数字档案室建设内容，对公司档案现行业务、管理流程、规章制度进行全面清理和剖析、梳理、整合、修订和完善，建立起科学而规范的档案管理制度体系，包括工作规章、管理制度、操作规范，涵盖公司级、分支机构级、各门类三方面。

### 6. 安全保密体系建设目标

从制度规范、设备、技术等多方面建立企业数字档案室安全保密体系，确保档案实体与信息安全。

## 三、建设内容

企业数字档案室建设任务主要包括应用系统建设、基础设施建设、数字资源建设、标准规范体系建设和安全保障体系建设等五个方面。

## 1. 应用系统建设

根据项目需求进行应用系统的开发建设,准备齐全基础数据,扩展对接系统范围(包括集中营运平台、大投行项目管理平台、统一风控平台、资产托管综合管理平台、固定收益审批事务平台等),做好接口规范,实现电子文件电子化归档,并通过逐步建设基础虚拟智能库房、异地异质灾备体系等。使其具备在档案门类管理功能、接收采集功能、分类编目功能、鉴定统计功能、系统管理功能等基础上,建立数字档案室专题数据库、不断优化大数据分析(见图 12-2)、人脸识别(见图 12-3)、签名比对(见图 12-4)等人工智能功能,不断进行档案资源挖掘、深化档案利用。通过实践,不断总结完善数字档案室的具体建设,提高系统性能要求,确保系统各项功能日趋成熟。



图 12-2 客户全档案生命周期图谱

## 2. 基础设施建设

在公司现有系统硬件的基础上,通过配备系统所需的网络基础设施、服务器、存储设备及必要的相关系统、软件等配套资源,逐步完成系统基础设施建设。数字档案室机房符合《电子信息系统机房设计规范》(GB 50174—2008)要求的 B 级机房标准。数字档案室所需的办公自动化设施和设备符合《企业数字档案馆(室)建设指南》的要求。



图 12-3 人脸识别查询



图 12-4 签名比对查询

### 3. 数字资源建设

建设应用系统所需的各种数据库，包括目录、全文、元数据库等。扩展对接的业务系统包括：集中营运平台、统一风控平台、资产托管综合管理平台、固定收益审批事务平台、大投行项目管理平台等。实现相应电子文件直接电子化归档，实现增量档案的电子化，存量档案的数字化。

### 4. 标准规范体系建设

根据电子档案管理实际情况，并参照国家、地方相关标准规范，采标为主，

新建为辅。建设适用于公司数字档案室所需的标准规范，保障对档案信息资源收集、存储、管理、共享、备份等环节的标准化和规范化，做到理论和实践并行。首先，修订《公司档案管理办法》《分支机构档案管理办法》，进一步明确公司全系统的档案管理体制和管控模式；其次，修订《电子档案管理实施细则》《文书档案归档文件整理规则》《客户资料档案管理实施细则》等不同载体和类型档案的具体管理措施；最后，编制《档案分类方案》《档号编制规则》《文件归档范围和保管期限表》《元数据方案》《数字档案资源备份管理制度》《数字档案室应用系统运维和安全管理制制度》《数字档案室突发事件应急预案》等操作规范和标准。

## 5. 安全保障体系建设

通过采用相应安全保障技术方法、配备必要的软硬件设施和建立健全数字档案室安全管理制度，确保数字档案数据的安全和应用系统及其网络平台的安全。健全电子档案安全管理制度，并通过逐步建设基础虚拟智能库房、异地异质灾备体系等不断优化档案资源安全保障体系。

# 四、建设技术方案

## 1. 总体架构

数字档案室建设是一项系统工程，根据《企业数字档案馆（室）建设指南》及《数字档案室建设评价指标》等相关要求，数字档案室建设应包括应用系统建设、基础设施建设、数字资源建设、标准规范体系建设和安全保障体系建设等。经过充分考虑各种业务及系统需求分析，综合考虑各种影响因素，形成数字档案室的总体架构，具体如图 12-5 所示。

数字档案室的总体架构由基础设施层、数据层、应用层、接入层共四层，以及标准规范体系、安全保障体系两大体系共同组成。

### （1）基础设施层

基础设施层是数字档案室的支撑平台，包括机房、网络、主机、存储、系统软件等信息系统运行的物理场所以及信息资源建设的工作环境，是任何一个信息系统正常运行所必需的基础设施。



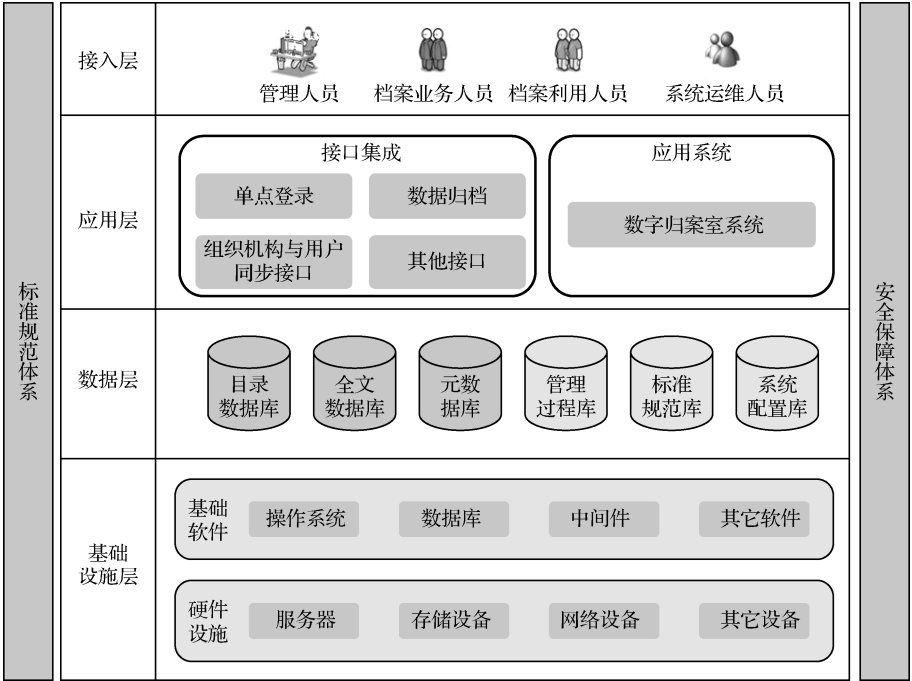


图 12-5 系统总体架构

（2）数据层

数据层是数字档案室的核心，包括各类电子档案信息资源和系统运行、管理、维护支撑资源两个组成部分，数据层一般由目录数据库、全文数据库、元数据库和管理过程库、标准规范库、系统配置库等组成。

（3）应用层

应用层是数字档案室建设的主体，包括接口集成以及应用系统，接口集成主要实现应用系统与其他业务管理系统的整合，确保应用系统成为信息化的有机组成部分；应用系统主要实现各类档案收集、管理、保存和利用的各项业务功能。

（4）接入层

接入层是指数字档案室的服务对象，包括管理人员、档案业务人员、档案利用人员和系统运维人员。

（5）标准规范体系

标准规范体系是数字档案室建设的指导方针，也是数字档案室发挥各项功能的基础，包含指导原则类、档案业务类、信息技术类和综合管理类的工作规

章、管理制度、业务规范。

#### (6) 安全保障体系

安全保障体系是数字档案室安全运行的保障,包括保障物理安全、网络安全、主机安全、应用安全、数据安全的一系列技术方案以及相应的安全管理要求。

### 2. 技术路线

(1) 参照 IS014721 开放档案信息系统(OAIS)模型,构建系统的应用架构。

(2) 软件平台架构。系统采用 B/S 架构;服务器操作系统为 Windows Server 2012 64 位;数据库系统采用 Oracle 11g;中间件采用 Tomcat;客户端操作系统为 Windows XP、7、10 或后续版本,软件应兼容各种版本的 IE 或非 IE 内核浏览器。

(3) 以 J2EE 为核心、基于构件技术搭建、基于 XML 的共享交换、基于 Web Service 的应用系统整合等技术路线。

(4) 系统设计要重视用户的使用体验,用户界面友好、设计简洁大方、功能齐备、美观实用、提示准确、响应快速,用户在无需代码或脚本编程的情况下,实现表单和流程等的客户化设计,能够满足系统易用性、易管理性的需要。

(5) 遵循国家、行业、国泰君安在信息技术、档案业务等方面的法规、标准、规范,能够满足数字档案室规范的建设、运行。

### 3. 系统部署方式

#### (1) 大集中式部署

##### ① 技术方案设计

根据国泰君安档案管理应用的实际需求进行大集中式部署。按照先进性、高可靠性、易维护性、高性价比的原则进行设计和部署。在实际部署上采用总部建立一套系统供集团内所有用户访问的方式。该方案遵循先进性、安全性、可靠性、灵活性等原则。

##### ② 技术方案特点分析

信息集中共享。由于系统所有用户单位共用同一套系统,各种信息和文档的共享变得非常方便,任何单位任何个人都可以通过授权方便地获得相关的信息。集团内部共享效率很高。

快速集中部署：由于系统主要设备都部署在集团总部，集团成员单位对网络等方面的要求相对较低，可以在集团部署后迅速部署。

维护方便：由于只在总部部署一套系统，所以容易建设，同时系统维护比较方便。

### （2）同城双中心

为保证档案业务的可持续性和数据的恢复性，建立同城双中心（唐镇、金桥），即为电子档案管理系统建立生产中心和灾备中心。然后根据情况选择磁盘数据的同步/异步复制技术，进行生产中心与灾备中心之间的数据备份传输，实现同城的灾难备份恢复，从而有效地管理风险、保证业务的连续运行，提高业务服务水平。

同城双中心主要由服务器和存储备份系统两部分构成。在生产中心配备两套数据库服务器来保证业务访问的稳定、高性能、快速响应及高可用性，而响应的数据则存储在生产中心的一套企业级存储系统上。同时，此系统通过磁盘数据的同步/异步复制技术将数据复制到同城灾备中心的存储系统上，生产中心与灾备中心的两套磁盘存储系统间建立磁盘镜像复制关系从而实现高可用性，保证关键数据的可恢复性与业务应用的可持续性。

## 五、建设进度

2016年，公司启动企业数字档案室项目，经过两年多的建设，项目一期、二期内容已完成，目前正处于三期建设中。项目一期完成了原档案管理系统常用功能开发，实现了多种类型档案的全生命周期，为数字档案室建设打下坚实的基础。二期主要实现了增量电子文件电子化处理，通过扩大电子文件接收范围及开发电子文件（电子档案）“四性”检测功能，将OA系统管理类电子文件、综合理财管理平台（非现场开户）电子文件等纳入其中，并引用大数据分析、人脸识别、签名比对、自然语音识别等技术进行档案资源挖掘、深化档案利用，入选了国家电子文件归档和电子档案管理首批试点单位。

项目三期是在一二期的基础上，建立统一接口平台，进一步规范扩展公司各业务系统电子化归档连接，实现与集中营运平台、大投行项目管理平台等对接，实现相应电子文件直接电子化归档，优化人脸识别、语音识别等人工智能技术应用，实现音频类档案智能识别、检索，建立ETL数据管理平台，提供电

子档案信息服务和支持,通过搭建虚拟智能库房和异地灾备系统,实现档案实体资源和信息资源的安全保管。建设进度安排如下:

2018年1—3月,项目启动实施阶段,组建项目团队,完成需求调研分析,修订完善制度标准规范,完成硬件采购与机房基础设施建设,并完成数字档案室系统建设规划与实施方案;

2018年4—9月,项目系统开发及平台建设阶段,完成硬件设备配置到位,档案资源准备齐备,数字档案室系统开发设计,实现传统载体档案数字化和电子文件电子化归档,以及档案资源深化利用和安全保密体系建设;

2018年9—10月,系统试运行及项目验收阶段,实现电子文件归档全过程管理和电子档案开发利用;

2018年11月,归纳总结试点工作文档,争取顺利通过全国企业数字档案室试点示范单位验收评审。

## 第三节 东风公司数字档案馆建设项目案例

### 一、项目综述

#### 1. 建设目标

按照《企业数字档案馆(室)建设指南》要求,通过基础设施建设、电子档案管理系统实施、电子文件归档和传统载体档案数字化、标准规范体系建设、安全保密体系建设等工作,实现档案资源数字化、档案管理自动化、档案传输网络化、利用服务知识化,实现企业档案工作升级换代、创新发展与提质增效,全面提升档案信息开发利用与共享服务能力,促进企业提高管理水平,增强核心竞争力,为东风汽车集团有限公司(以下简称“东风公司”)持续健康发展提供有力支撑。

#### 2. 项目规划

项目一期(2016—2017年):搭建集团总部级数字档案馆,完成对总馆、

各分馆及各直管单位实物档案的数字化处理并实现集中存储管理。

项目二期（2018—2020 年）：建立集中式档案信息资源管理平台。实现与各业务系统有效集成，完成电子档案的自动收集，归档，安全保管和整合利用。

项目三期（2020 年之后）：搭建集团档案目录共享中心，实现集团总部档案管理系统与下属单位独立部署档案管理系统中目录数据的集成共享。

目前已完成项目一期建设，正在筹备项目二期。

## 二、项目一期建设过程

项目一期共分八个阶段进行。

### 1. 预研究阶段

主要工作是研究项目可行性，形成预研究报告，并召开评审会，邀请有关领导专家对报告进行评审。

### 2. 基本设计阶段

主要工作是编制基本设计书、业务需求清单、系统功能清单、画面清单、票据清单、接口清单等。

### 3. 商务阶段

主要工作是产品选型、商务招投标，形成产品选型测试报告、商务合同、技术协议书等文档。

### 4. 内外部设计阶段

主要工作是供应商进场与公司档案部门、信息部门共同完成业务要件、功能定义书、页面流转图、接口定义书、数据库设计书的研编工作。

### 5. 开发实施阶段

主要工作是供应商技术团队、数字化扫描团队按项目计划开展相关工作，

并向公司信息部门提供系统源代码。

## 6. 综合测试阶段

主要工作是开展综合测试，编制综合测试报告。

## 7. 业务测试阶段

主要工作是开展业务测试，编制业务测试报告。

## 8. 移转上线阶段

主要工作是编制移转上线材料，上会评审。

# 三、一期项目建设效果

## 1. 项目完成情况

### (1) 基础设施建设

以信息系统安全保密要求为建设目标，完成机房、网络设施、硬件设备、基础软件、其他终端设备的搭建。

电子档案管理系统部署在东风公司私有云平台，并将数据库安装到物理机上，以此来保障数据安全，具体见图 12-6。

### (2) 电子档案管理系统建设

按照公司档案工作相关标准，建设具有收集、整理、保存、统计、利用、系统管理、传统载体档案辅助管理等功能的电子档案管理系统。

### (3) 数字档案资源库建设

#### ① 传统载体档案数字化

完成 130 万页馆藏档案数字化，近 800 件荣誉实物档案、1666 (A1) 张图纸数字化，形成可在线利用的数字化档案。

#### ② 系统集成

定制开发接口 5 个，实现公文系统文件归档集成、统一用户管理系统集成、用户登录统一认证集成、消息中心集成、温湿度采集集成等。

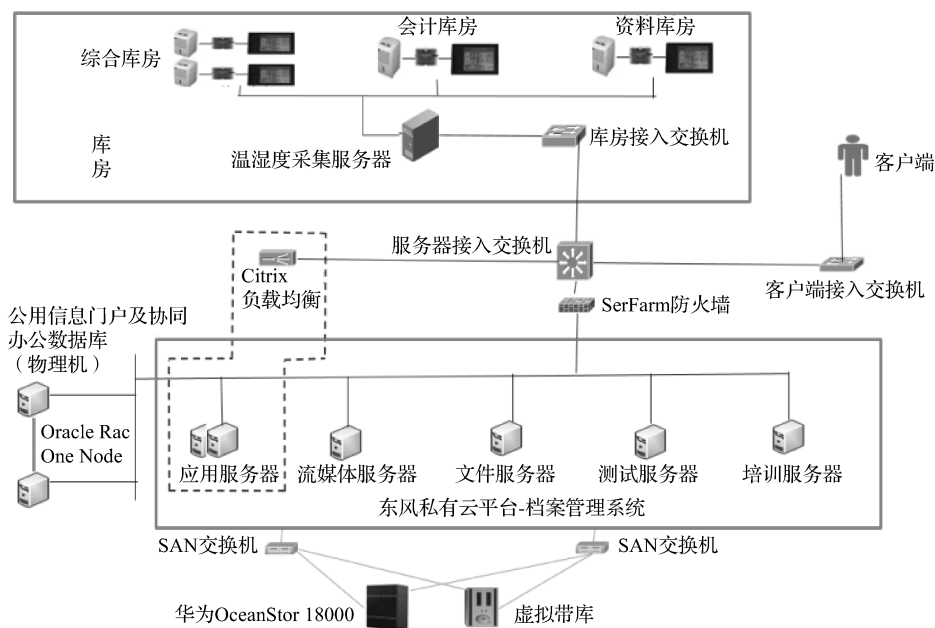


图 12-6 东风公司私有云平台架构图

### ③ 电子档案长期保存库建设

为实现电子档案安全存储，系统对不符合长期保存要求的电子文件进行格式转换，使其满足电子档案长期保存的需要。

### (4) 制度规范建设

项目实施过程中产生项目管理文档 67 篇，形成知识积累，对公司其他信息化建设项目提供借鉴参考。项目管理文档分布情况如图 12-7 所示。

### (5) 安全保密体系建设

① 系统安全采用 2 级安全防护，并依据东风公司总部系统 2 级安全防护规范进行部署。

② 系统安全分为硬件要求（Infra）、软件要求（开发）、业务要求（客户）。

③ 东风公司总部采用防攻击、防病毒、防篡改功能、漏扫系统、安全审计等安全防护措施，并设置安全工程师应对安全策略。

④ 客户端通过 SSLVPN 方式访问。

⑤ 东风公司总部采用物理安全措施，包括规范机房管理流程、安排专职人员对机房环境进行监控等。

系统安全保密体系部署情况如图 12-8 所示。

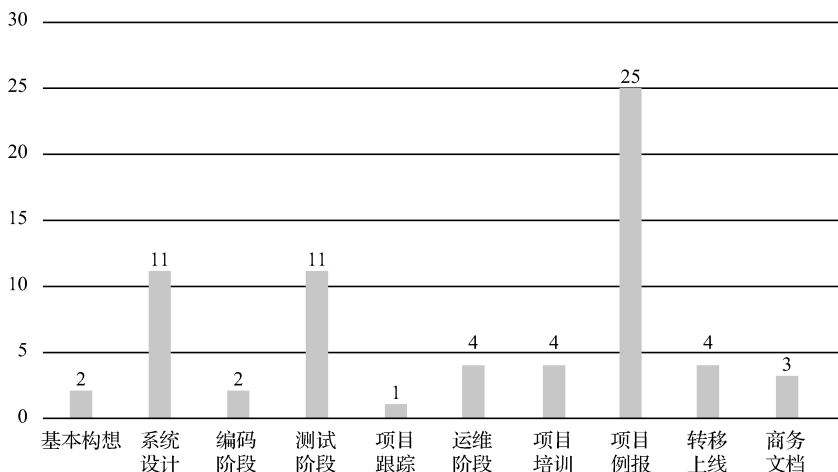


图 12-7 项目管理文档分布情况图

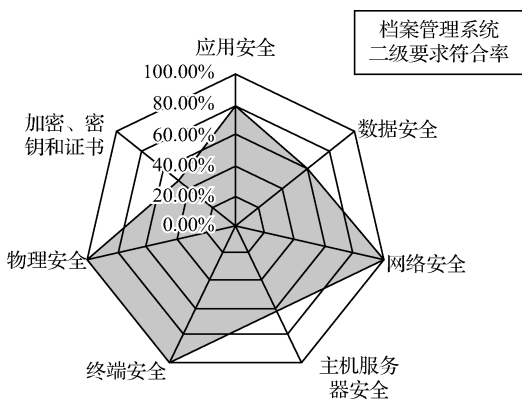


图 12-8 安全保密体系部署图

## 2. 开展项目建设效果评价工作

至 2017 年年底，一期项目已上线运行近半年，档案部门与信息部门共同组织人员对数字档案馆一期项目进行了效果评价，全面评估一期项目建设成果，并以此作为二期项目预研究的依据和条件。评价工作从交付文档、系统评价、收益评价、使用评价、专家评价等五方面开展。评价比重分为系统评价权重 35%、ROI/KPI 收益的评价权重 35%和使用评价权重 30%。

### (1) 交付文档

重点考察项目基本构想、投资立项报告、基本设计书、安全定级、移转判



断报告、竣工单、建设效果等。

(2) 系统评价

设置评价标准，从项目计划完成度、系统功能实现度、性能参数是否满足设计要求、用户体验满意度、应用架构设计水平、遗留问题改善、运行维护等 7 个方面设置评价点，一一对照、逐项打分。

(3) 收益评价

从 ROI、KPI 两项分别考察系统是否达到预期效果。

(4) 使用评价

通过访谈和查看系统数据，了解用户对系统功能的使用情况，计算用户使用功能与系统全部功能的比例（以服务合同中的功能为评价参考标准）；从运维方考察系统数据维护的准确性等。

通过自评，公司总部数字档案馆一期项目获评 85 分。

四、项目建设经验总结

1. 健全的项目组织体系是保障数字档案馆项目进程的前提

公司成立了数字档案馆项目指导委员会、项目领导小组、项目管理办公室，分级负责、层层把关，采用 RACI 矩阵方式进行项目监管（具体方案见图 12-9），采用里程碑的方式把控整个项目进度（具体方案见图 12-10）。

项目RACI矩阵						
	项目领导小组	项目管理办公室	业务组	应用架构组	基础架构组	量子伟业
预研究	A	A/R	R/C	R		
基本设计	A	A/R	C	R	R	
商务阶段	A	A/R	C	R		
外部设计	A	A/R	C	C	R	R
( 内部设计 )	A	A/R	C	C		R
开发实施	A	A/R	I	R/C		R
综合测试	A	A/R	I	C		R
业务测试	A	A/R	R	R/C		R
移转上线	A	A/R	I	C		R

R = Responsible, 执行人 A = Accountable, 批准人  
C = Consulted, 咨询人 I =Informed, 被告知人

图 12-9 项目 RACI 矩阵

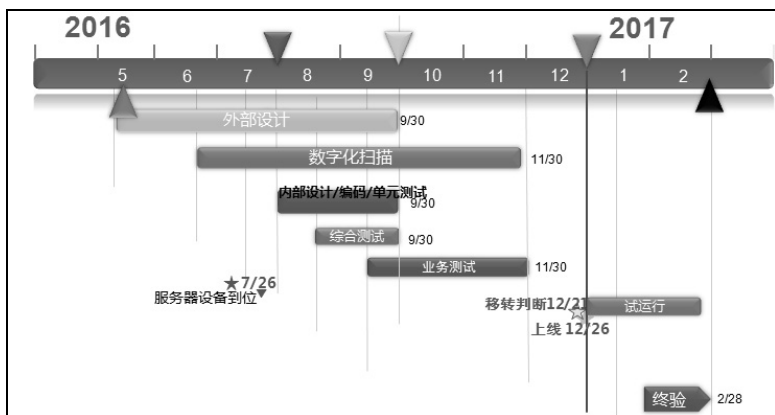


图 12-10 项目进度里程碑

### （1）指导委员会职责

- ① 负责掌控整体方向，指导合作原则和战略合作意向的制定。
- ② 负责裁决需求顺序冲突，指导解决合作过程中的重大问题。
- ③ 负责项目阶段里程碑评审决策。

### （2）项目领导小组职责

负责关键业务决策，保障项目资源，项目阶段里程碑评审等。

### （3）项目管理办公室职责

- ① 负责制定项目计划及资源安排。
- ② 负责协调各小组工作进度的一致性。
- ③ 负责协同各项目关系人的项目问题及实现目标，定期召开项目例会或专题会议，并向项目领导小组进行阶段汇报。

## 2. 清晰的业务需求和技术需求是奠定项目建设的基石

（1）以业务需求为导向，精心梳理档案工作流程，分析档案业务现状，提出业务改进目标，明确业务需求、软件功能要求和项目建设构想。

（2）档案部门、信息中心、供应商三方紧密合作开展需求调研。

- ① 制定调研计划，组织调研团队。
- ② 对总馆、分馆和二级单位进行实地调研。
- ③ 形成《业务需求调研报告》。

（3）在业务需求调研、技术需求调研的基础上，经充分论证，确定系统设

计方案。

- ① 根据《业务需求调研报告》进行需求设计，形成《用户需求设计书》。
- ② 根据《用户需求设计书》形成《系统功能需求表》。
- ③ 结合集团信息化建设情况，形成档案管理系统和第三方系统接口需求。
- ④ 统计增量和存量数据，评估档案资源库存储规模。
- ⑤ 根据系统注册用户数量计算服务器 tpmc 参数，评估软硬件配置需求。

3. 多轮次的充分测试是项目后期良好运行的保障

(1) 产品选型测试

通过 5 轮应用验证性测试（POC 测试）和产品试用，充分了解供应商产品性能，验证其产品与项目建设需求的满足度，具体见图 12-11。

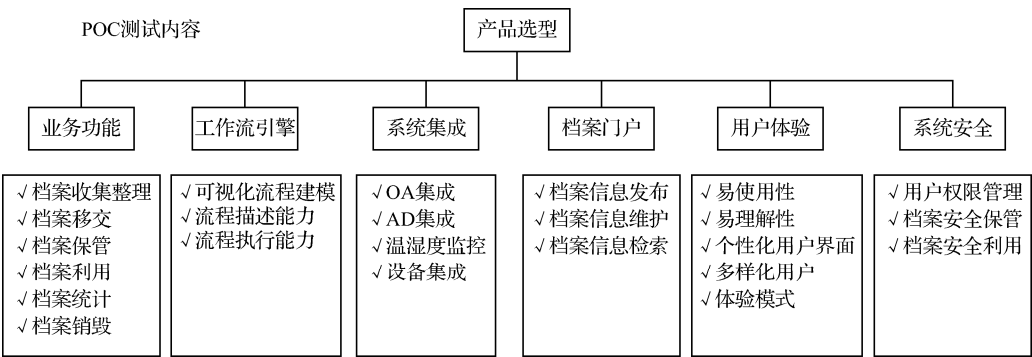


图 12-11 POC 测试内容图

(2) 项目验收并试运行

4. 良好的沟通和管理是项目建设的必要条件

(1) 良好的项目质量管理沟通

- ① 宣贯项目实施方法、文档规范。
- ② 评审项目交付物质量。
- ③ 评审项目过程文档。
- ④ 测试系统功能。

(2) 良好的项目进度管理

- ① 评审各阶段计划（WBS）。
- ② 跟踪项目进度。
- （3）良好的项目风险管理和课题管理
  - ① 识别项目风险，上报项目课题。
  - ② 编制项目课题清单，进行项目课题管理。
- （4）良好的组织协调管理
  - ① 协调人员、办公地点、会议室、网络、硬件设施等各类资源。
  - ② 定期组织专题会议。

#### 5. 健全的运维体系是项目有效应用的有力保证

以公司信息中心为主体组成数字档案馆管理系统运维团队，主要职责是负责系统运行维护、数据备份和处理业务需求变更等工作。

## 第四节 本钢集团有限公司数字档案馆建设案例

### 一、系统建设概况

本钢集团有限公司（以下简称“本钢集团”）档案信息系统是以电子文件归档管理为主、传统载体档案数字化采集上传为辅的数字档案管理系统。历经一期、二期工程逐步建成完善。

一期工程于 2011 年 6 月启动。实施范围和目标，是开发集成具有收集、整理、保管、利用等基本功能的数字档案管理系统，开发 ERP、OA 系统数据归档接口，实现 ERP 系统（含采购、储运、生产、质量、销售、设备、财务、人资、工程管理等子系统）、OA 系统等业务系统电子文件的在线自动归档以及馆藏档案数据的采集录入，实现企业档案信息资源数字化、网络化利用共享及规范的安全管理。历经需求分析、归档范围及流程模板梳理、取数逻辑分析配置、接口程序开发集成及功能测试、用户操作培训、数据读取检测、系统集中压力测试、归档文件鉴定著录维护等主要阶段。2012 年 9 月 27 日正式上线运行。

二期工程于 2013 年 8 月启动，实施范围包括档案信息系统异地灾备应用接

管建设、B/S 架构功能优化、ERP 电子文件归档扩展、设备图纸数字化管理试点。2014 年 11 月 25 日正式上线运行。

## 二、系统运行管理体制

### 1. 系统应用范围

系统应用范围包括本钢集团所有部门、子公司及其下属立档单位。截止 2017 年年底，系统共设置 144 个全宗节点；432 个 OA 归档节点（涵盖了本钢集团 OA 系统所有有归档价值的收文、发文、请示报告、部门业务表单等）；归档 ERP（采购、销售、生产、质量、工程、设备、人资、财务等八大专业子系统）电子文件表单 364 种。是一个动态接续归档、实时在线利用的数字档案信息系统。

### 2. 系统运行管理体制

系统运行采用“统一归档、分级负责、集中管理”的管理体制。

“统一归档”即本钢集团统一建设档案信息管理系统，对 ERP 系统、OA 系统等信息系统电子文件进行统一归档。

“分级负责”即本钢集团各部门、各单位 OA 系统文书员、模块员及 ERP 系统各主管部门业务人员，负责各部门、单位应归档电子文件的检验、提交或上传工作；各立档单位档案员负责本单位 OA 系统归档电子文件的鉴定、著录、归档维护、利用服务等工作。

“集中管理”即本钢集团对档案信息系统服务器实施集中配置，对归档数据进行集中存储和安全备份，对超权限利用实行集中审批管控。

### 3. 管理职责

#### （1）集团档案馆职责

主要负责用户权限管理；组织机构树、档案分类树、档案资源树创建维护；用户线上业务进度质量监督指导；系统用户培训、咨询指导；线下档案资源数字化转换上传；在线利用审查与推送服务；ERP 系统归档数据及档案馆直接负责的三个档案全宗 OA 电子文件鉴定、著录、归档维护等。

## （2）集团信息化中心职责

主要负责系统功能模块优化完善、扩充；系统运行维护、故障处置；OA系统、ERP系统归档电子文件定期读取归档、数据备份；系统用户培训、咨询指导等。

## 4. 服务器配置情况

配置应用服务器、热备服务器、中间服务器、异地灾备服务器等四台服务器。应用服务器、热备服务器、中间服务器部署在本钢集团信息中心机房，异地灾备服务器部署在本钢集团能控中心机房。其中应用级灾备在本地集群端两台服务器共享6T的磁盘阵列，在异地灾备中心部署容灾服务器并增加6T的磁盘阵列，实现数据及应用的安全备份。

本地服务器集群与异地灾备服务器间保持数据库同步、电子文件同步、应用同步。在本地集群出现因网络中断、机房断电、存储故障等异常情况下，异地灾备服务器可在5分钟内实现应用接管。

# 三、科学设计电子文件归档技术方案

为满足现代信息技术条件下信息利用批量、集成、快速响应等特点，档案项目部组织ERP、OA系统各专业主管部门认真进行需求调研分析，全面开展电子文件归档范围、展现表单、取数逻辑的梳理确认。

遵循国家档案局关于电子文件原文信息、元数据及背景信息齐全的要求，将电子文件流转全过程形成的有价值信息全面纳入归档归集范围。借鉴、采纳实施公司比较成熟的理念、流程，根据本钢集团业务流程和管理结构，开发集成符合企业业务流程的多种检索模式，以满足各方面工作的查考利用需求。

## 1. OA系统电子文件归档技术方案

将有保管价值的发文、请示、报告、收文、部门业务、知识管理、电子期刊、体系认证等内容全部纳入归档范围。将电子原文、附件、原始稿、修改痕迹、主表单等形成完整的业务链条，满足内容查检和过程追溯的利用需求。

## 2. ERP 电子文件归档技术方案

### (1) 归档方式

根据 ERP 电子文件数据集成展现的不同情况,分别采取不同的归档、归集方式:对在 ERP 系统中有集成固化的单据凭证和报表展现样式的,通过接口程序自动读取数据合成电子文件实现归档;对无固化展现样式但内容有长期利用价值的,通过接口程序归集相关的目录信息以供查询;对财务、人资系统有关难以提供取数逻辑但有归档需求的 ERP 报表,采取部门定期上传归档。

按照 ERP 各业务流程的特点,将 ERP 有关专业系统中产生的相关业务单据建立业务链式检索,实现整个业务过程产生的业务单据的完整追溯。

通过对 ERP 系统中的工程项目主数据、设备主数据的归档,形成工程项目档案和设备项目档案的管理目录和资源存储结构,利用该结构实现各种相关资源的归集化检索利用。

### (2) 检索方式

主要采用业务链式检索模式。

#### ① 采购凭证链

将 ERP 采购系统、物流系统、财务系统产生的业务单据相互追溯,形成链式查询。根据不同的采购类别进一步细化出材料采购、原料采购、设备备件采购、小原料采购、废钢采购、零库存采购等凭证链。

#### ② 销售凭证链

将 ERP 销售系统、质量系统、财务系统中的业务单据相互追溯,分别形成国内销售凭证链和销售出口凭证链。

#### ③ 设备分层检索

本钢集团 ERP 设备树结构为:集团—厂矿—生产线—区域—设备—设备部件(零件)。结合设备编码规则,利用来自 ERP 设备管理系统的设备主数据,构建设备树结构,设计管理配置平台,实现对设备资源的灵活配置与管理。其中设备层及以下层次都可以虚拟归集档案资源,设备主数据及归集资源随 ERP 系统定时归档而动态增加或更新。

设备分层检索资源库可归集的信息主要包括:部件维修基准清单、部件维修点检清单、工作任务单、设备档案中的施工文件、竣工图纸等。

#### ④ 项目树检索

本钢集团 ERP 工程项目树结构为:公司代码—年度—项目类型—项目—项

目投资控制分解结构。

结合工程项目编码规则,以工程项目主数据为基础,建立工程项目树结构。项目主数据及归集资源随 ERP 系统定时归档而动态增加或更新。

项目树检索资源库可归集的信息主要包括:项目合同明细目录、工程投资情况汇总表、OA 系统流转的工程项目审批文件及基建档案中的施工文件、竣工图纸等。

#### ⑤ ERP 会计账检索

设定统一的查询检索界面,检索 ERP 财务系统会计账,根据检索到的目录信息可对应查询检索记账凭证。

通用查询条件可以设定公司别、会计科目、周期、会计类别、户号、参号、辅助项、传票日期等。

#### ⑥ ERP 报表检索

归档的 ERP 报表,根据不同的业务系统分开查询检索。

主要分为财务报表、工程报表、原料系统报表、质量系统报表、生产系统报表、人资系统报表、设备系统报表、销售系统报表、上传报表等。

通过以上系统功能的实现,使档案信息系统的信息归集能力和服务能力得到了保证。

### 3. 全面、细致地开展电子档案完整性、真实性、可用性检测确认

信息化建设向来追求“三分建设七分管理十分数据”,档案信息系统建设更是必须全面细致地开展归档数据检测确认工作,其主要原因,一是易因 OA 系统公文流转结束或 ERP 系统业务办结导致专业人员忽视数据检测工作。档案系统归档的 OA、ERP 归档数据均是业务流程已完结的数据,专业人员此时已不再过多地关注其真实性、完整性问题,加上对归档数据的利用也仍可通过现行的 OA 系统、ERP 系统实现,由此极易导致相关用户对归档数据的真实完整及长远利用重视不够。二是易出现因逻辑分析不准而导致归档数据不准。

ERP 数据归档的电子文件是通过整合后的取数逻辑关系进行模板化处理实现的,数据归档前期可能存在部分因逻辑关系不严密而导致归档数据与 ERP 业务系统数据存在差异的情况。

#### (1) 检测确认方式

为了做好本钢集团档案信息系统已归档数据的完整性、真实性、可用性检



测，档案馆、信息化中心组织实行专业部门、基层单位用户检测与档案馆系统管理员全面检测相结合，以档案馆系统管理员检测确认意见为主。

一是采取 ERP 各专业模块主管业务人员、OA 系统文书员、模块员等在历史数据正式归档上线前全面检测、确认、备案；日常每批次归档当期数据时进行检验、确认、提交。

二是采取档案信息系统管理员获得临时授权，同时登录 ERP 系统、档案系统检查核对同一电子文件的数据对应情况。

## （2）检测确认内容

### ① OA 系统

检测电子原文是否可以正常打开；原文标题与内容是否关联一致；原文内容是否准确；文件正文、附件、原始稿、修改痕迹、主表单是否齐全；数字印章及领导电子签名是否正常显示等。

### ② ERP 系统

检测档案系统中数据与 ERP 系统中数据是否完全一致，主要包括：同一单据在 ERP 中有正式数据在档案系统中也应有正式数据；同一时间段的数据在档案系统与在 ERP 系统中应数量相等；同一种类型的单据在档案系统中与在 ERP 系统中的账套情况应相符；同一条数据在档案系统中与在 ERP 系统中内容应完全相同。

## 四、建设成效

本钢集团档案信息系统实现了对集团公司重要电子档案信息的集中安全管理，提高了档案信息传输利用效率，降低了企业文档管理成本，推进了本钢集团档案业务建设标准化、规范化进程，也大大缓解了 OA 和 ERP 两大业务系统的运行压力。

截止 2017 年年底，本钢集团档案信息系统共归档 ERP、OA 等信息系统电子文件 2296 万余件。其中 ERP 电子文件 454.4 万件、ERP 目录表 1801.6 万条、文书类电子文件 37.4 万件、工程合同审批电子文件 2.1 万件；共上传数字化设备图纸 6831 张；上传照片、荣誉等类电子图片 7950 张；上传各类纸质档案目录 34 万余条。档案馆编制的职工调转录用人索引、大事记、组织机构沿革、本钢年鉴等编研成果也上传至系统首页，方便了各部门、单位的查询利用。

2014-2017 年，系统电子档案在线利用 51619 人次计 97402 件。

2014 年 11 月，国家档案局组织召开全国企业档案信息化建设现场会，本钢集团应邀发言，介绍档案信息化建设案例。

2015 年 8 月，本钢集团档案管理信息系统课题被评为“2015 年辽宁省企业管理进步成果二等奖”。

2016 年 12 月，本钢集团档案管理信息系统课题被国家档案局、国家发改委列为“企业电子文件归档和电子档案管理”试点项目，目前试点工作在有序推进中。

2018 年 3 月，本钢集团档案管理信息系统课题被辽宁省档案局列为“企业数字档案馆（室）建设试点”项目。

## 五、主要经验

### 1. 选好数字档案馆建设范围和突破口

遵循国家档案局“以需求为导向，统筹规划，循序渐进，选准突破点，关注投资效用”等原则要求，档案馆会同企业信息化部等专业部门，对国家有关规定及外部各类档案机构的信息化案例进行了充分调研、借鉴。先后实地调研观摩了上海宝钢、天津石化、辽宁电力、辽宁移动等档案信息化建设先行单位不同实施公司的典型案例，强化了对本钢集团档案信息化实施范围、分项推进步骤、技术路线的认识。

通过开展全面的案例调研分析，认为本钢集团信息系统五级架构第四级的 ERP、OA 等信息系统已运行稳定成熟，其强大的信息集成能力使得电子文件的成果固化、内容展现已具备了电子档案的特点。档案信息化建设的突破点，应放在这些增量、利用率高、领导和管理人员关注且系统安全风险较大的电子文件归档管理上，把 ERP、OA 系统中的存量、增量信息资源首先管好用好。为此确定开发实施范围为：开发 ERP、OA 系统数据归档接口，集成具有“收集、整理、保管、利用”基本功能的数字档案管理系统，实现 ERP、OA 等业务系统电子文件的在线自动归档以及馆藏档案数据的采集录入，实现数字化、网络化利用共享及规范的安全管理。

## 2. 加强档案、信息技术、专业部门及基层企业协同

面对本钢集团信息系统多业务系统、多帐套的技术管理架构，归档数据量大，实现有归档价值电子文件的全面归档和有序规范管理，面临着许多技术和管理难题，绝非档案系统一家单打独斗就能胜任。为此，集团公司成立了由主管副总经理、档案、信息化、相关业务部门及主要子公司负责人参加的项目管理组织架构；档案馆、信息化部分别指定专人任项目主管和系统管理员，相关业务部门指定专人全过程参加项目建设管理。

建立了工作例会制度，负责协调和落实项目开发集成、数据检测等相关问题。

加强对内组织协调，档案项目部组织 OA 系统及 ERP 系统八大专业主管部门（采购储运、销售、财务、生产、质量、人资、设备、工程）全面进行功能需求分析，梳理确认归档文件范围、业务流程及业务链集成需求，共同开展功能需求分析、归档数据分析配置、中间检验归档平台开发集成、用户权限分类配置、用户培训、系统功能集成配置、全面测试整改等重点环节工作。

在系统集中压力测试阶段，各相关部门、基层单位组织各自用户在固定时段全面上线检验利用系统各项功能，及时反馈发现的疑难问题。

在系统招标评标、数据切换、测试运行、验收上线、异地灾备接管测试等关键节点，本钢集团信息化部系统组、技术组等专家均参与查验把关，强化后台技术支持，二期工程中信息化中心软件开发人员还直接参与项目开发的具体工作，避免了档案部门孤军奋战的知识短板。

## 3. 加强项目进度、质量管理考核

### （1）加强对内考核点评工作

2011 年 6 月档案一期项目启动后，档案信息系统即纳入本钢集团信息化考评范围。档案馆会同信息化中心充分利用信息化例会点评、纪要指导、线上巡查、系统监控、线下示范指导、用户全面培训、电子邮件催办等形式，对用户检验提交、鉴定、著录编目、归档维护等环节工作进行全面、全过程监督指导，推动了电子文件数据读取、检验提交、价值鉴定、著录编目、归档维护工作逐步实现了常态化、规范化。

### （2）加强对外协调联络

档案馆、信息化中心及时全面地向实施公司介绍本钢集团 OA、ERP 系统

的技术架构、业务流程，一同研究档案信息系统开发集成的目标要求，及时提供各项电子文件的取数逻辑，配合进行归档数据分析配置工作。

档案馆、信息化中心负责人与实施公司总部高层保持热线联系，遇有现场项目经理难以解决的问题，及时向实施公司高层提出整改建议并充实力量。

通过周、月例会等形式定期分析汇总存在的问题，双方及时达成整改共识，从而使项目的进度、质量处于可控和有序推进状态，确保了功能需求分析、归档数据分析配置、中间检验归档平台开发集成、用户权限分类配置、用户培训、系统功能集成配置、归档数据完整性真实性可用性检测等节点目标顺利实现。

### （3）加强线上抽查与质量进度监管

档案馆、信息化中心加强对项目运行的监督指导。每月均统计、发提醒电子邮件或电话催办，保证了归档渠道畅通、归档节奏紧凑，保证了“统一归档、分级负责、集中管理”体制有效实施。

## 参考文献

- [1] (澳) Leszek A Maciaszek. 需求分析与系统设计[M]. 金芝, 译. 北京: 机械工业出版社, 2006
- [2] 王英, 蔡盈芳, 黄磊. 电子文件管理[M]. 北京: 清华大学出版社, 2016
- [3] 胡思康. 软件工程基础(第2版)[M]. 北京: 清华大学出版社, 2015
- [4] 仇伟, 郭洪远. 云计算的特点和应用展望[J]. 数字技术与应用, 2011(04): 168-168
- [5] 程妍妍. 云计算环境下国外政府电子文件管理面临的挑战与实践[J]. 浙江档案, 2016, (2): 14-17
- [6] 牛力, 韩小汀. 云计算环境下的档案信息资源整合与服务模式研究[J]. 档案学研究, 2013, (5): 26-29
- [7] 张怀廉, 季雪岗, 李慎民. 企业数字档案馆的建立(之一): 江苏省电力公司数字档案馆概况[J]. 中国档案, 2005, (3): 58-59
- [8] (美) 大卫·芬雷布. 大数据云图[M]. 盛杨燕, 译. 浙江: 浙江人民出版社, 2014
- [9] 陈萍, 张涛, 赵敏. 信息系统安全[M]. 北京: 清华大学出版社, 2016
- [10] 全国档案工作标准化技术委员会. 档案工作标准汇编 9[M]. 北京: 中国档案出版社, 2010
- [11] 全国档案工作标准化技术委员会. 档案工作标准汇编 10[M]. 北京: 中国标准出版社, 2015
- [12] 蔡盈芳. 企业数字档案馆结构研究[J]. 档案学研究, 2008, (1): 26-30
- [13] 蔡盈芳. 航空产品知识管理关键技术研究[J]. 航空科学技术, 2012, (6): 50-54
- [14] 蔡盈芳. 具有多层属性实例知识检索研究[J]. 图书情报工作, 2012, (23): 108-112, 134
- [15] 蔡盈芳. 企业档案信息的知识管理研究[J]. 档案学研究, 2014, (2): 42-46

- [16] 蔡盈芳. 基于本体的航空产品知识库构建研究[D]. 北京: 北京交通大学, 2011
- [17] IDC. 中国公有云市场报告[EB / OL]. [http: //blog. sina. com. cn/ s/blog\\_750050490102wt8u. html](http://blog.sina.com.cn/s/blog_750050490102wt8u.html), 2015
- [18] 崔海莉, 张惠达. 云计算环境下档案信息管理系统风险分析[J]. 档案学研究, 2013, (1): 56-60

# 后 记

写一本有关企业数字档案馆建设的书已经酝酿很久了。早在 2014 年，鉴于企业档案信息化建设的需求，笔者就有了这样的想法，并着手收集相关资料。当时，企业数字档案馆建设尚未进入高潮，提法较多的是档案信息化，因此，选定的题目是“企业档案信息化建设理论与实践”。近几年，随着形势的变化和档案事业的发展，数字档案馆建设成为企业档案工作的重要内容。考虑到数字档案馆建设与档案信息化内容基本相同，并结合当前企业档案工作的重点，遂将本书定位为企业数字档案馆建设。随着国家档案局对企业数字档案馆建设工作的加强，笔者觉得更有必要出一本系统介绍企业数字档案馆建设的书，供广大企业档案工作者参考。本书撰写历经几年，一方面是在参阅了大量有关文献基础上撰写而成，另一方面也是笔者工作经验的总结。

在书稿付梓之际，首先，要感谢各级领导对写作本书的悉心指导；其次，感谢广大中央企业档案工作者给予我了解需求和深入学习企业数字档案馆的机会；第三，要感谢提供案例的企业及其档案部门领导，这几个案例对充实全书内容有着重要意义；第四，感谢业内专家，在书稿形成过程中，我大量借鉴并参考了他们的文献，有的已列出在参考书目中，有的因各种原因没有列出；最后，要感谢在书稿形成过程中帮助我校对、审核的亲人和朋友，正是有了他们的大力支持和帮助，使书稿得以完成！

衷诚希望本书能为广大档案工作者提供借鉴、参考和启发。

蔡盈芳

2018 年 8 月于北京

